Designing and Evaluating an Advanced Dance Video Comprehension Tool with In-situ Move Identification Capabilities

Saad Hassan
Tulane University
New Orleans, LA, USA
saadhassan@tulane.edu

Caluã de Lacerda Pataca
Rochester Institute of Technology
Rochester, NY, USA
cd4610@rit.edu

Laleh Nourian
Rochester Institute of Technology
Rochester, NY, USA
ln2293@rit.edu

Garreth W. Tigwell
Rochester Institute of Technology
Rochester, New York, USA
garreth.w.tigwell@rit.edu

Briana Davis
Rochester Institute of Technology
Rochester, New York, USA
bhd8713@rit.edu

Will Zhenya Silver Wagman
Tulane University
New Orleans, Louisiana, USA
wsilverwagman@tulane.edu

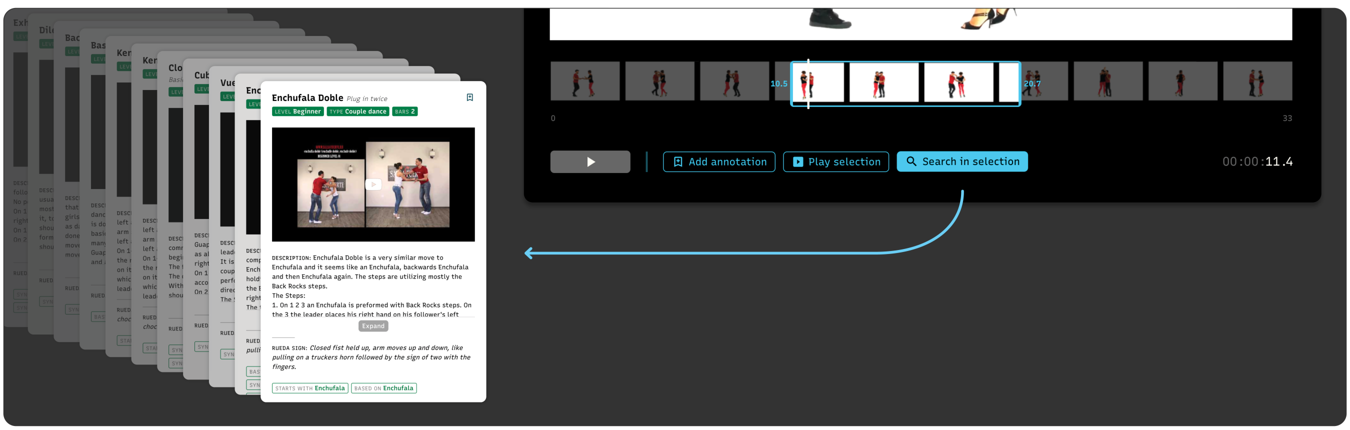


Figure 1: In this paper, we present a dance video comprehension tool with span-based auto-identification functionality.

ABSTRACT

Analyzing dance moves and routines is a foundational step in learning dance. Videos are often utilized at this step, and advancements in machine learning, particularly in human-movement recognition, could further assist dance learners. We developed and evaluated a Wizard-of-Oz prototype of a video comprehension tool that offers automatic in-situ dance move identification functionality. Our system design was informed by an interview study involving 12 dancers to understand the challenges they face when trying to comprehend complex dance videos and taking notes. Subsequently, we conducted a within-subject study with 8 Cuban salsa dancers to identify the benefits of our system compared to an existing traditional feature-based search system. We found that the quality of notes taken by participants improved when using our tool, and they reported a lower workload. Based on participants' interactions with our system, we offer recommendations on how an AI-powered span-search feature can enhance dance video comprehension tools.

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642710

CCS CONCEPTS

• **Human-centered computing** → **User studies**; *Laboratory experiments*; *Graphical user interfaces.*

KEYWORDS

Dance, Dance Learning, Choreography, Human Movement Analysis, Search Systems

ACM Reference Format:
Saad Hassan, Caluã de Lacerda Pataca, Laleh Nourian, Garreth W. Tigwell, Briana Davis, and Will Zhenya Silver Wagman. 2024. Designing and Evaluating an Advanced Dance Video Comprehension Tool with In-situ Move Identification Capabilities . In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3613904.3642710

1 INTRODUCTION

Dance is a popular art form [106] characterized by aesthetic and often richly symbolic body movements [25]. For instance, poses in classical or formal dance styles, such as Bharatanatyam [6, 30] or Hula [96], hold profound symbolic significance [114] and encapsulate cultural narratives. This holds true for some more prevalent dance styles in the U.S. as well, such as Salsa [39] and Ballet [109]. Since many dance styles incorporate intricate techniques, expressions, and formations, dance learners need to engage in imitation, repetition, mental simulation, practice, and extensive performance observations [93]. An important first step in dance learning is often the analysis of movements [94], which can be achieved through dance videos [3, 51, 86, 116].

Certain dance forms, such as ballet, sometimes adhere to a pre-established glossary of moves, prompting learners to focus on identifying sequences of named dance moves. Some digital resources for these forms offer comprehensive video catalogs that include formal names and metadata detailing associated properties like length, type, origin, body parts involved, and significance [39, 109]. However, the challenge arises when users seek to identify an unfamiliar dance move using digital video catalogs, as it necessitates specifying various properties. For individuals immersed in video-based dance learning, this involves stepping out of the context of video-watching and selecting filters to uncover the desired name or information about a particular dance move. This process may appear challenging and cumbersome for dance learners. *To streamline this experience, especially for dance styles that adhere to defined movement glossaries, we have developed and evaluated an advanced comprehension tool that features an automatic dance move identification functionality.*

First, we conducted an interview study with twelve dance learners to explore their previous experiences watching dance videos and note-taking. Through thematic analysis, we uncovered the behavioral insights and challenges that learners encounter during video analysis and note-taking tasks using currently available tools. Those findings inform the design of our video comprehension tool and the selection of stimuli for the subsequent study.

Notably, our tool enables users to automatically identify dance moves in a selected video segment. For this, we utilized the Wizard-of-Oz method, a well-established approach in HCI for exploring future technology [24, 41, 55, 88]. Despite considerable attention and advances within the computer vision community for the automated recognition of dance poses [20, 21, 38, 85, 107], including approaches with accuracy rates in the low 90's for specific dance styles [56, 63], improvements are still needed to perfect segmentation in dance pose estimation and recognition [27, 84]. This seems likely, given how recent years have seen promising advancements in gesture segmentation for complex motion sequences [57], and related areas like sign language [79]. As such, the Wizard-of-Oz method allowed us to focus on exploring interface paradigms that, in the future, may allow the integration of dance pose recognition into a notes-taking framework.

To do so, and following the first interview study, we conducted an experimental study with eight additional dance learners to study how dance learners engage with Cuban Salsa videos during a video analysis task using our system, and if dance pose recognition technology would effectively aid them in the process. Our findings revealed the benefits of our comprehension tool with in-situ dance move auto-identification capability over a feature-based identification approach, showcasing improved note quality and enhanced subjective performance, coupled with a reported reduction in workload. Observational findings revealed participants' interactions with the system, including using the span-selector and auto-identify features, and highlight remaining challenges.

Our findings inform computer vision researchers and designers of interactive technologies to support dance learners. To summarize our contributions:

- We provide empirical evidence demonstrating that a dance video comprehension tool, equipped with video-span-based auto-identification features for dance moves, enhances the user experience for dance learners. Further, we offer initial insights into the usage of this auto-identification system, particularly in understanding user-generated segments' frequency and accuracy.
- We identify a novel task for computer vision researchers: utilizing a user-defined segment of a formal dance video as input for a system that automatically identifies dance moves.
- Based on our interview findings and subsequent prototype design and evaluation, we propose design recommendations on how to employ a timeline-based video playhead and structure the results list in a dance video comprehension tool featuring in-situ move identification functionality.

2 BACKGROUND AND RELATED WORK

HCI researchers have been developing tools to help dancers conceive and produce dances. Since dance is a kinesthetic experience [103], many of these technologies have focused on providing visual idea generation that supports the iterative and interactive nature of the choreographic process [17]. While some researchers have focused on using augmented reality [54, 61, 62, 97], virtual reality and motion capture technology [22, 62, 65], as well as 3D videos for learning dance [91, 92], our paper is centered on tools to support *video-based dance learning*.

2.1 Using Videos for Dance Learning

Like several other fields involving structured human movement, e.g., sports and exercises, videos for learning dance are also becoming increasingly popular [10, 20, 42, 51]. Video-based learning in e-learning "Dance and Choreography" programs has been encouraged [122], and learning from observation (visual learning) can also lead to better learning outcomes for recall compared to verbal instruction alone [74]. In recent years, especially since the COVID-19 pandemic, there has been an increased pedagogical shift in dance learning from in-class learning to remote learning, which involves greater use of recorded videos [74]. It has been shown that students benefit from online peer feedback on their dance recordings through comments and ratings [52], and activities such as collectively watching dance videos through Zoom's share screen feature have also become common [52, 74].

Beyond simply enhancing visual learning, the use of videos in dance education extends to appreciating performances, enabling peer feedback, and fostering a sense of community among learners. This is evident in the research of Alves [4] and Doughty et al. [33], which explores the dissemination, storage, and display of dance performances to address limited rehearsal time and space.

Cognitive scientists have also studied how dance learners process and understand videos. Research has found that for dance, the ability to chunk continuous action streams into ordered, distinct, and interpretable units plays a vital role in facilitating motor learning, and such learning does transfer across dance forms [31, 60]. A recent study further revealed high agreement among dance learners on the boundaries of individual dance steps [78]. This suggests that leveraging movement boundaries can benefit video analysis tools, as explored in our own study.

2.2 Describing and Visualizing Dance Moves

Capturing and documenting dance remains notoriously challenging. Dancers utilize diverse modalities to articulate and convey individual moves and sequences, including drawings, emotional descriptions, and time analyses [26]. Researchers and practitioners have further explored notational systems based on parameterizing body movement, shape, and spatial usage, e.g., using Laban Movement Analysis [66–68].

Researchers have also designed interactive screen-based visualization tools to capture and visualize dense dance movements using computer animation-based information objects and interactive graphics [43, 43, 83]. LifeForms, for instance, is a graphical interface that enables choreographers to sketch movement ideas in space and time. Famous American dancer and choreographer, Merce Cunningham, used it in his renowned dance "Trackers." The system involved constructing movement sequences by selecting stances from menus through chance procedures, determining involved body parts, the number of limbs moving simultaneously, and the type of movement [98]. LifeForms was later adapted into DanceForms, a choreography software that facilitates visual and kinesthetic learning, enhancing understanding of movement and choreographic design through on-screen visual representation of choreographic ideas [44].

Faced with the challenge of synthesizing dance moves using systems like the symbolic notation or the Laban description of body structure at a specific time [66, 67], other researchers have concentrated on enhancing keyframing-based motion with expressive elements, e.g., EMOTE Model [23] and other multidimensional motion interpolation approaches [95]. The EMOTE model blends key poses and Laban's Effort-Shape theory [68] to produce more naturally expressive synthetic gestures. Researchers working on computer-aided choreography have also developed dynamic models consisting of dance verbs, such as "to jump," "to flip," etc [53].

Given the inherent difficulty of describing dance moves and the variety of approaches, learners would greatly benefit from tools that provide access to exemplar videos aligned with corresponding segments in their own videos. Rich textual metadata that combines simplified terminology with named dance move videos could further empower learning and detailed note-taking for accurate recreation. These considerations prompted us to explore the textual and visual content of individual result snippets in our system.

2.3 Interactive Tools for Dance Learners

Research in dance education has highlighted the cognitive dimensions of technologically supported instruction, emphasizing the design principles inherent in creative multimedia technologies [28, 36]. Drawing inspiration from these creative design principles, researchers have proposed systems to support the dance community [2, 15, 18, 19, 32, 35, 90, 101, 112].

2.3.1 Computational Choreography Tools. For over five decades, researchers have investigated computational choreography tools and automatic composition for dance, primarily for expert dancers and choreographers [69]. Some of the earliest systems, like LifeForms, have evolved incrementally over thirty years, transitioning from stance-based to sequence-based timelines, enabling users to work at a higher level and edit individual figure timings separately. Despite these advancements, users have underscored the usability challenges of reshaping paradigms, emphasizing the need for flexible systems that accommodate pre-existing approaches [99].

Several other automatic composition tools have also been designed. Swarm Simulation Toolkit, another early system, allows for dance movements based on the provided aesthetic rules [119]. Similarly, COMPOSE offers an interactive environment, allowing users to simplify tasks by switching between spatial and temporal views. It also provides a realistic animation of the final result [17].

Researchers have also explored using genetic algorithms to generate choreography for ballroom dance [70, 71, 81]. These algorithms, comprised of defining a movement vocabulary, initializing movement sequences, generating mutants, and selecting mutant sequences, ultimately create a choreography [71]. Selection criteria are later used to test the "fitness" of the sequences [70]. Researchers have postulated that some of these approaches can be utilized for other ballroom dance forms with the same rigid structure but different types of steps and timing [81].

Researchers have explored automatic composition and simulation tools designed to serve as both support systems for dance teachers and independent study aids for students [12, 17]. These tools can generate dance sequences distinct from those produced by humans and have the potential to enhance the human creative process, encouraging choreographers to view art from diverse perspectives. Soga et al. [105] further contributed to the line of research by classifying basic ballet steps into meaningful categories and fragmentary families for automatic ballet composition, introducing the "Web3D Dance Composer" for choreography using a predefined step list [104, 105].

While these systems primarily target expert dancers and choreographers, our focus differs as we aim to assist non-expert dancers in learning and understanding. Our approach involves enabling them to deconstruct existing choreography videos into individual moves instead of creating them from scratch. Despite this distinction, we can draw insights from the extensive research in expert-oriented systems to inform and enhance our investigation and the design of our proposed solution.

2.3.2 Tools for Annotating Dance Videos. More relevant to our work are annotation tools for dance. A recent retrospective on two decades of dance research and HCI mentions several tools, including Choreographer's Notebooks [121]. These tools provide multi-modal annotation capabilities, allowing choreographers and dancers to annotate video segments of dance videos remotely and asynchronously using various input forms, such as written comments, sketches, images, and video demonstrations.

One of the earliest annotation tools introduced a Tablet-based tool that facilitates multi-modal video annotation for contemporary dance videos [14]. This tool allows users to incorporate voice and pen-based (stylus) annotations on recorded videos. The ongoing work aimed to leverage motion tracking to define the dynamic behavior of annotations. They also proposed capabilities to embed images, web pages, and other videos within their annotations. Similar work introduced a choreographer's notebook system that allows users to add comments on the video timeline in the form of color-coded text and sketches [101]. The system's interface also allowed users to maintain a journal documenting notes and progress. Certain choreographers' notebooks have also adopted specific annotation schemes to describe the stage positioning of dancers in videos [112]. Other systems offer users expanded functionalities, including annotation marks, audio descriptions, grouping comments [15], and peer-feedback activities [3]. However, these annotation tools were developed before recent advancements in dance move recognition. As a result, none of them have explored the potential of these advancements to enable new features, including in-situ movement/pose identification.

El Raheb et al. designed an archival system offering browsing, searching, visualization, personalization, and textual annotation for dance recordings (motion capture data, video, and audio). Their primary focus was to support dance education by providing accessible and intuitive tools. Notably, they employed a user-centered, iterative design approach, ensuring the system met the needs of its intended audience [36]. Their evaluation focused on four key features: searching and browsing by genre and metadata, locating specific recordings and annotations, using the timeline, and adding personalized annotations. Similarly, Singh et al. explored the design of commenting features for dance videos in a choreographer's notebook, specifically focusing on insertion, exploration, and editing functionalities [102]. These two papers provide design guidelines in systems analogous to ours, closely aligning with the recommendations that emerged from our two studies.

2.3.3 Archival Systems with Dance Move Searching Functionality. Recently, researchers have introduced extensive archival systems comprising repositories of multimodal dance recordings, accompanied by associated motion capture, audio, and visual data, to support dance analysis and learning (e.g., [36, 72, 89]). These systems may also allow manual annotations to be added along the video timeline, creating benchmark datasets for machine learning research, enriching dance video collections with expert insights, and stimulating discussions within the dance community through shared movement analysis. Some systems, e.g., BalOnSe, facilitate advanced searches using domain-specific vocabulary and metadata such as titles and featured dancers (e.g., [37]). While these systems are valuable for the dance community and research, they are not necessarily designed for learners to look up unfamiliar dance moves.

Prior work on retrieval in the context of performance arts has also shown that dance learners and viewers would like to search for dance moves and sequences that exhibit certain affective qualities expressed by the dancer, and performance details, e.g., history, country of origin, dance costumes worn, spatiotemporal information about dancers, etc. [58, 59]. The extensive list of properties underscores the critical importance of properly contextualizing dance styles for effective understanding, even if the specific properties may vary depending on the use case.

Commercial repositories of pre-labeled isolated dance moves, often called dictionaries, also exist online for various dance styles, such as Ballet [109] and Salsa [39], and allow feature-based searches. While these dictionaries are informative, they still require dance learners to know the dance move names or a substantial amount of other information for retrieval. For dance learners, a system that can does not require them to possess the name of the dance move would be more beneficial.

Advancements in machine learning may soon enable the creation of systems where users submit a video of a dance move to identify its name and retrieve performance details [29, 76, 110, 115, 120]. In fact, some recent approaches have shown movement recognition accuracy in the low 90s when applied to dance videos downloaded from YouTube, i.e., recordings with inconsistent quality [63]. While these technologies are currently under development and improvement, a need exists to explore their potential utility for dance learners, how they could be integrated into video analysis tools, and how the user experience of video-based search function should be structured. Through interviews, we gathered firsthand insights into the experience of watching dance videos and taking notes as part of the dance learning process, informing the design of our system.

3 INTERVIEW METHOD

There has been limited prior research on the experience of dance learners engaged in the educational activity of watching a video to learn a new dance, especially regarding their note-taking behavior and identification of named dance moves. To this end, we asked:

> **RQ1:** What are the experiences and challenges of dance learners who watch videos to write notes?

To answer this research question, we conducted interviews to investigate the difficulties that dance learners experience when closely watching dance videos and trying to take notes.

3.1 Interview Procedure

3.1.1 Interview Participants. To recruit participants, we posted an advertisement across various university dance groups and shared it on a dance subreddit. The advertisement included two questions: "Are you currently involved in dance or choreography?" and "Have you been engaged in dance or choreography in the past 5 years?" We invited individuals who answered yes to at least one question to participate in the study. Based on the standard interview participant sample size for similarly scoped studies [16], we recruited 12 dance learners, 11 women, and one non-binary person. The median age of participants was 21 (σ=2.6). Table 1 provides more information about the participants, including their dance experience and the styles that they specialize in.

3.1.2 Protocol. Participants chose the location of the interview, either on our university campus (n=2), in a dance studio (n=2), or over Zoom (n=8). If the participants preferred Zoom, we asked them to conduct the interview in a room with ample space for them to move freely in case they would like to demonstrate a dance move. Participants signed an IRB-approved consent form for participation at the start of the study. Participants' responses to demographic questions were also obtained at the start of the study. The interviews lasted an average of 42.4 minutes. Participants were paid $40 for participation in the interviews.

Participant	Age	Dance Experience	Styles
D1	19	More than 3 years	Ballet, Salsa, Jazz, Tap
D2	20	Between 1 and 2 years	Ballet, Contemporary, Salsa
D3	21	Between 1 and 2 years	Hip hop
D4	21	Between 2 and 3 years	Ballet, Jazz, Tap
D5	21	More than 3 years	Kathak, Ballet, Lyrical Hip Hop, Salsa
D6	20	More than 3 years	Salsa, Tap, Jazz, Hip hop, Contemporary
D7	20	More than 3 years	Ballet, Jazz, Hip Hop, Tap, Palm
D8	18	More than 3 years	Salsa, Ballet, Modern, Lyrical, Irish, Jazz, Tap
D9	21	More than 3 years	Salsa, Jazz, Ballet, Hip Hop, Tap, Kick
D10	21	Less than a year	Lyrical, Musical Theatre
D11	21	More than 3 years	Tap, Ballet, Hip hop
D12	29	More than 3 years	Salsa, Bachata

Table 1: Age, dance experience, and dance styles of participants recruited in the interview study.

3.1.3 Interview Guide. In the first half of the interview, we asked the participants to share their dancing experiences using the story interview technique often used in prior HCI literature on dance [113]. The participants were asked to describe their journey of learning to dance, and the interviewer took notes while the participant freely expressed their experiences. During this process, if needed, the interviewer asked follow-up questions to gain a better understanding of the participant's dance learning journey, including the styles they engaged in, when they started, their level of experience, any formal dance education they received, their performance history, and the frequency of their dancing activities.

In the second half of the interview, we focused on comprehending the participants' approach to learning dance, specifically using videos. We asked them to elaborate on the entire process, which varied depending on the participant. We inquired about the techniques they used to memorize and recall dance sequences. We investigated participants' note-taking practices. We inquired about the mediums they use for note-taking, the nature of the content, their methods for referencing different segments of a performance, and their preferred structures for organizing notes. To contextualize our questions, we referred back to the participant's prior dance experiences shared in the first half of the interview.

3.2 Interview Analysis

We video-recorded and transcribed each interview, assigning a unique participant code (D1-D12) to every dancer. Our transcripts also included detailed descriptions of any dance steps demonstrated and corresponding video time stamps. To analyze the data, we conducted a reflexive thematic analysis combining deductive and inductive approaches [13]. Two members of our team watched the interview recordings, read through all the transcripts, and generated codes to identify important themes. Throughout the coding process, we remained open to identifying bottom-up themes.

4 INTERVIEW FINDINGS

All twelve dancers mentioned watching dance videos. Seven dancers mentioned using YouTube, six mentioned watching short-form social media dance clips, and four mentioned different television shows. D8 described their experience: *"I watch a lot of social media dance accounts, so I do get inspiration from there definitely... visual learning is the biggest thing... it's just hard to explain without physically watching a performer's body doing it."* D8 also stated that even expert choreographers watch dance tutorial videos as it allows them to learn alternative ways to teach: *"...if they're watching the tutorials, it can help you find different ways to explain the movements to other people that might make more sense."*

During our interviews, five dancers reported that their focus while watching dance videos extended beyond simply observing the sequence of steps. D2 expressed this sentiment, stating: *"I'd be looking at like the moves but I'd also be like looking at the dancer,... their emotions, [and the] message. Are they just having fun and showing off their talents, or is there, something deeper going on. What is the environment? What's the background?"*

D6 further emphasized that at times it is important for them to understand the deeper meaning behind a performance, stating: *"Some contemporaries are really sad and have a really intense story and then others are just like beautiful and light-hearted and that sort so... I'm watching with like a critical eye, but also watching for the sake of watching."* Other dancers noted that they watch videos that have a *shock factor* (D5), *leave them in awe* (D3, D7), or *have sharp moves* (D8). Despite a strong interest in watching dance videos, dance learners often find it challenging to analyze these videos effectively when using them for learning purposes.

4.1 Challenges with Video-based Dance Learning and Current Workarounds

Participants described the challenges that they experienced with video-based dance learning. Remembering the sequence of dance moves in a video was found to be a challenging task by eight dancers. According to D1, *"Some people are able to just watch and memorize, but some actually have to do the moves to memorize the dance."* Additionally, the difficulty of memorization can depend on the genre of dance as stated by D1: *"In tap, there are really quick little movements, whereas for Jazz, it's a lot of bigger movements, but it's not as fast-paced as tap. So for me, I can just watch and memorize a Jazz dance."* D8 added that even though they find it easier to memorize individual moves, it can be more challenging to remember the transitional steps that connect the more notable moves: *"It usually takes* *me a little bit longer to connect those different segments together... the little in-between is where it can get a little more difficult to do."*

Five dancers mentioned that they were accustomed to practicing in studios with their teachers facing in the same direction as they are. However, online videos are often mirrored, and the dancer faces the camera instead. D6 found this to be an issue, stating: *"I find it impossible to learn when someone records the person by facing them, whereas like when you record from behind them, it makes more sense because, their right arm matches your right arm..."*

Dancers also described the challenge of learning from faster videos, with some being sped up to look more impressive. D4 commented: *"Sometimes they're like sped up so it's definitely harder to learn something when it's sped up to look nice."*

Participants also mentioned their **existing workarounds** when they had difficulty following dance videos. Seven dancers in our study reported frequently pausing while watching dance videos to focus on specific snippets, as stated by D5: *"I'll usually watch other people's full choreography and then, just pause on like interesting snippets."* Dancers also noted the benefit of watching videos at a slower speed, gradually increasing the speed as they learn. D5 acknowledged the usefulness of slowing down and how tutorial videos on YouTube already slow down the dance moves, stating: *"I would slow it to 0.5 speed or something, and try to follow it. Other times, there's more formal tutorials where like the choreographer will actually talk to you through every single move slowly."*

Seven dancers reported that they typically **learn dance video segments in parts** rather than as a whole. D1 elaborated on this method, stating, *"Instead of teaching the whole thing in one go, they'll teach you little segments so that it's easier to remember and you kind of just keep building on top of it."* Dancers described that this approach facilitates easier memorization of complex dance sequences by breaking them down into smaller, more manageable parts.

D8 further explained that dance moves are often broken down into individual components and then combined using *"tiny movements that combine portions throughout the piece."* Four dancers mentioned that even within individual moves, sequences of movements are often broken down into three to nine steps, depending on the style. Different dancers have different methods of identifying these steps. For instance, D9 explained that *"you have people that count music 1-2-3-4-5-6-7-8, people that make noises [gave some examples of sounds], and people who mix noises and numbers."* D5 who specializes in Kathak mentioned that *"the counts have like specific enunciation, so instead of saying like 1-2-3-4, it would be bha-tat-thei-tat."* They further mentioned that these step count identifiers may be commonly known within the dance community or might have been developed in a specific studio.

Overall, our participants reported employing strategies that involve leveraging video player features and breaking dance routines in videos down into smaller segments or steps to streamline the analysis and learning process. Despite utilizing these workarounds, memorizing a dance routine from a video is a non-trivial task. Consequently, video analysis during dance learning is often complemented by extensive note-taking.

4.2 Note-taking as a Method for Learning Dance

For dancers aiming to pursue a professional career in a formal dance, it is important to accompany their video analyses with written notes that contain correct labels and precise enunciations for identifying the sub-components of a routine. This practice not only aids in memorizing complex dance moves but also facilitates learning technical terminology associated with various moves and communicating it with others. D5 mentions the significance of this approach, stating: *"[...] the enunciated counts... and certain elements... like a spin with your hands... have specific names... our professor insisted that we jot those names down to have them memorized, in case another dance teacher inquires about them in the future."*

D6 said that note-taking and journaling were encouraged by teachers on recorded performances: *"Our teachers would always encourage us to have journals and like to write down your corrections."* D8 stated that note-taking involves writing down your thought process for some moves: *"Writing down what you did helps you remember where you're going and where your thought process was at that time. So you can kind of pick up where you left off."* They further mentioned that it would be useful to attach notes to specific segments of video: *"...if you're trying to jog your memory of the choreography, having notes in there specifically attached to each part of the video would be really helpful tips to remember"*

Participants also talked about the challenge of textually describing a dance move while taking notes: *"It's a lot of almost like reverse engineering, you know you're watching them and trying to sit you're looking at it like okay how in the world do it (describe it)."* D7 expressed a preference for using common terminology in their notes: *"I'd probably honestly use terms that are more common in athletics, in general,.. a sumo squat you probably know, but a better idea of what I'm talking about."*

Four participants also emphasized the importance of understanding formations in understanding dance videos and transitional steps and mentioned their note-taking practices for both:

> D8: *"It was a mix of the names of the dance moves and then stuff in my own words, that wouldn't really make sense to anyone else. And then another thing... was draw the formation that we were in. So they would put Xs on a piece of paper, and then label which one was whom... So you kind of remember where you were as well."*

In summary, our participants described how complementing video analyses with detailed written notes is beneficial. They emphasized the utility of personal notation systems, including the use of common terminology and other representations, to document their analyses. Textually articulating individual moves poses a challenge for dance learners, which motivates participants to learn the formal names of dance moves.

4.3 Identifying and Naming Isolated Dance Moves

Five dancers, each specializing in at least one dance form that follows a pre-established glossary, expressed interest in learning the names of formal dance moves. D8 described using online tutorials for ballet, saying, *"So everything in dance [ballet] has a name. So I would usually just go on YouTube, look up the name of the trick, and then they'll usually have tutorial videos if it's nothing too obscure,* *they usually have tutorial videos. And those are really helpful just because they kind of go through step by step exactly what you do..."*

D8 highlighted that in dance styles like Ballet and Salsa, move names often carry meaning. Specifically, she mentioned:

> D8: *"There have been countless amount of times I've taken written notes, which seems like it wouldn't be the case. But in Ballet specifically, it is one of the most important parts is the names of the moves and they're all French... One of the moves is called Pas de Chat... which means step of the cat... it makes more sense because the moves have a name for a reason. It correlates to me, helping remember moves when learning."*

A minority of our participants mentioned not relying on formal analysis. D6 explained their method that involves references to different body parts involved: *"I think if I was explaining it to someone I would try to go through like each part of the body and what it's doing. For tap [moves] like for a Paradiddle [performs a scuffle, followed by step heel, all on the same foot] it's just your feet... if I were to describe with you, it would be heel-toe-toe-heel all on the same foot."* D9 pointed out that even when they learn about a dance move's name, there are often regional differences in its naming: *"People in different regions call it different things. So, what's called a Calypso in one place might be called a Barrel-Turn in another... Having a better understanding of terminology would be helpful for people who are trying to learn."*

When participants knew the name of the dance move, they mentioned using search and hashtags to look for performances of unfamiliar moves, especially on short-form social media. D1 commented: *"Sometimes even like Instagram is like the big help because there are like, they can like to search up using the hashtags"* Participants also discussed what they liked to see in the results. For example, D11 mentioned that it is nice to have the descriptions in a step-by-step manner: *"depending on how long the dance section is and so you'd break it up by you show someone a whole eight count."*. D10 mentioned that descriptions online are often too long and it would be better if they were more succinct: *"[...] like short and concise it usually is pretty good because sometimes YouTube ones are long."* Other dancers had other suggestions for descriptions, e.g., keywords (D4), meta-data (D11), and cultural information (D10).

Most participants noted the difficulty in identifying and describing dance moves for note-taking without knowing their formal names, a challenge that extends to searching for dance moves as well. Some suggested technological solutions to alleviate the challenge of looking up unfamiliar dance moves online. D6 envisioned a tool similar to Shazam for movement-based identification: *"Shazam the song... like pick out what song it is. The tempo of the song can help you find it. You could use movement as a similar tempo if it could identify the tempo for you, that can be really helpful."*

Overall, our participants highlighted that learning the formal names of dance moves is useful in dance learning, as it streamlines the process of identification of dance moves within a routine and improves the quality of notes. This notion conceptually influenced our system design. With the rapid progression in dance move recognition technology [20, 21, 38, 56, 63], there is an opportunity to further assist participants in finding unfamiliar dance moves without leaving the video analysis and note-taking context.

5 EXPERIMENTAL STUDY METHOD

Our interview study revealed valuable insights into dancers' experiences with video-watching, note-taking, and comprehension. This revealed a clear need for a system that helps them create high-quality notes while learning dance moves, especially for styles with established formal move names. These findings informed the design of a prototype system aimed at streamlining this note-taking and dance move lookup process.

We evaluated this prototype in Study 2. To enable a comparison with a suitable baseline, we designed and adapted our system for Cuban Salsa, a specific dance style with an existing glossary of moves [39]. Using these two prototype systems, we sought to answer our second research question:

> **RQ2:** In a comparison between the experience of users who performed a series of dance video annotation tasks using a video comprehension tool with dance move identification capabilities versus an online catalog of dance moves:
> (a) Is there a difference between task performance and workload?
> (b) How did they make use of the systems' features?
> (c) What challenges did they still face?

5.1 Prototype System Design

Findings from the interviews informed the design of our dance comprehension tool in multiple ways. The interface is shown in Figure 2, and it is described in detail below. A video figure illustrating different features of the prototype is also available as supplementary material.

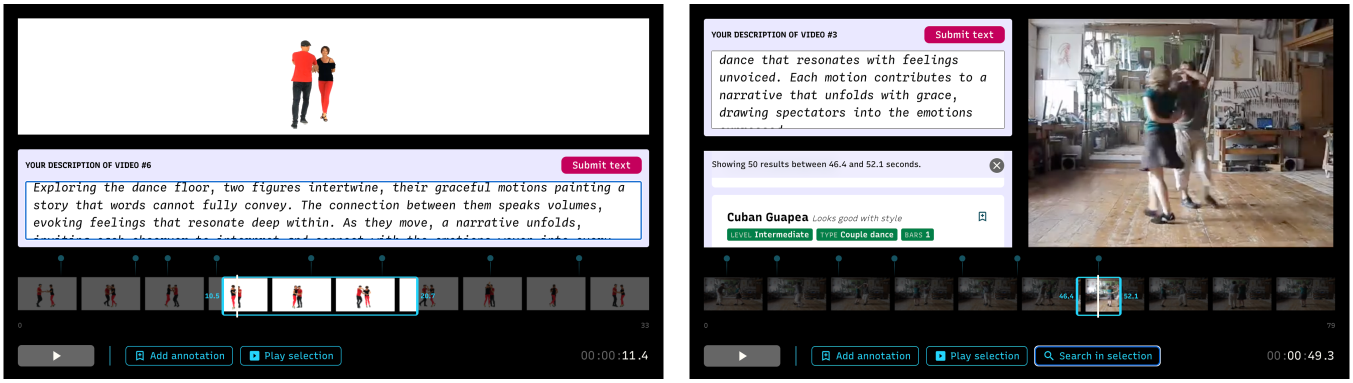
5.1.1 *Two ways of playing.*

Participants in the interview study discussed competing needs for watching dance videos. Some aimed for a broader perspective, looking into the dancer's motives, emotions, and message. Others focused on finer details: discovering intriguing moments, re-watching specific parts, and toggling play-pause intermittently.

Our devised solution is a timeline interface that hopes to accommodate both requirements. It incorporates a play-pause button that controls a play head that, as in a traditional video player, will traverse the video from start to finish (Figure 3a). Complementing it, the timeline includes a span selection functionality, allowing users to confine the play-pause dynamics to a specified segment of the video (Figure 3b). To help adjust the boundaries of the span selection, the timeline is illustrated with sequential screen captures extracted from the video, serving as visual reference points.

5.1.2 *Note-taking.*

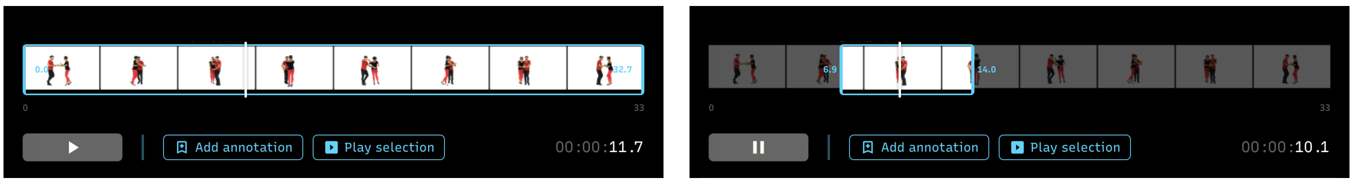
Participants expressed a desire to associate notes with specific segments of the video. In the prototype, they would have the option to create annotations linked to specific time spans within the video. These annotations appeared as turquoise-colored pins on the timeline. When these pins were selected, their corresponding video segment was highlighted. This feature is demonstrated in Figure 4. Additionally, the interface also included a larger, free-form note-taking component, a response to participants' comments on their needs for a tool that also allows for journaling, e.g., writing down thought processes, corrections, and so on.



(a) Interface with span-selector, timeline annotation, and note-taking features, but no dance move auto-identification feature.

(b) Same as Figure 2a, but with added auto-identify functionality.

Figure 2: Comparison between how the note-taking interface was displayed between the two conditions.



(a) Initial state of the timeline

(b) Span selection from 6.9 to 14.0 seconds in the video.

Figure 3: Wizard-of-Oz prototype timeline controls allowed participants two options: playing the entire video using the gray button, or selecting a portion with the turquoise-colored *Play selection* button. When dragged, the play head's movement across the timeline remained independent of span limits, but when the video was playing it adhered to whichever of the two buttons were used.

5.1.3 Auto-identification Feature. Besides note-taking functionalities, interviewees showed an interest in being able to identify and learn details about dance moves they would encounter while watching a video, such as their formal names, detailed descriptions, meta tags, and so on. In section 5.2, we describe our Wizard-of-Oz approach to simulating a computer-vision search within the videos, but in terms of the auto-identify functionality, as it was presented to participants, the prototype had a *Search in Selection* button that returned results of dance moves found *within* the current video span. These results were listed to the left of the video and included the move's name, both in its original language and an English translation; standardized meta-data descriptions, as requested by some participants, e.g., step counts, difficulty levels, body parts involved, etc.; a video example of that one move—including the move shown in different angles, also something participants had requested; a detailed description of it; among others. See Figure 5 for an example.

5.2 Wizard-of-Oz Protocol

With Study 2, our aim was to examine participants' utilization of integrated auto-identify capabilities during a dance video understanding and note-taking task (RQ2).

While advancements in computer vision for dance pose recognition and movement tracking for specific dance styles have been made [20, 21, 38, 85, 107], the technology is not yet at a stage where to allow a fully functional recognition system as envisioned by the needs of participants in Study 1. Recent research in dance recognition [56, 63] and related fields [57, 79] suggests these capabilities might be imminent, at least for some dance genres. However, for our purposes, we opted to employ a Wizard-of-Oz protocol, where the dance moves were manually annotated before the experiment. This allowed us to concentrate on the observational aspect of the study without being hindered by the limitations of current technology.

As mentioned earlier, to enhance the empirical investigation of the tool's usage, we tailored our system to a specific formal style: *Cuban Salsa*. We chose it due to its well-defined list of dance moves [39] and the keen interest shown by participants in our interview study who were professionally acquainted with Salsa. Among the six selected videos, five were categorized as advanced or higher in terms of difficulty by the largest available online glossary of Cuban Salsa moves [39]. The remaining video was classified as "Improver" (Intermediate) in difficulty level, yet it features more than two dancers, which adds the challenge of describing formations and different moves performed by four dancers. The description of video stimuli is provided in the Appendix A.

We selected videos with at least ten dance moves for a more comprehensive investigation into the functionality of the "Play

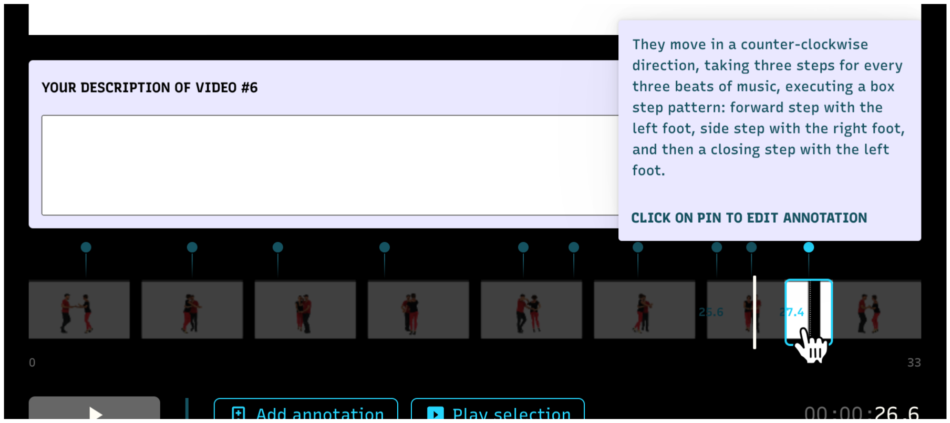


Figure 4: Example of an annotation added to the timeline and highlighted by the user. Note that there are many other faded-turquoise-colored pins attached to different positions on the timeline.

selection" and "Search Selection" features. This served to discourage participants from searching the entire video span – given our Wizard-of-Oz protocol, this would likely result in a large number of dance moves appearing in the results list.

To ensure consistent time and effort investment in note writing across the two experimental conditions, we maintained a minimal difference of no more than two dance moves between the videos. Each participant viewed three videos across both conditions. We also made sure that the video quality was satisfactory, ensuring that the dance moves were visible and unobstructed. Importantly, to solely rely on participants' prior knowledge or our tool for note-taking, the videos employed in the study deliberately lacked annotations or voice-over descriptions that could aid in identifying the dance move names.

The compilation of movements displayed in the results stemmed from a pre-processing of each video, as follows:

(1) A team member watched these annotated videos to identify the sequencing and starting and ending time stamps of each movement. These annotations were then confirmed by a Salsa instructor. The Salsa instructor had three years of teaching Salsa and more than ten years of performing.
(2) To replicate the potential outcomes of utilizing an actual automated identification system, a curated set of results was manually compiled for each movement. This simulation was designed to reflect the experience one might encounter in future real-world use. Concretely, for each movement, a selection process was employed that involved identifying the closest match and other dance moves that exhibited visual resemblance to the given movement. These selections were made from a compiled catalog of Cuban Salsa dance moves [39]. While compiling this "match list," priority was given to choosing movements that shared attributes with the designated movement. These attributes included body parts involved, number of steps, and duration. Furthermore, we included related dance moves from the existing catalog as well. This way, we simulated a plausibly imperfect machine-learning system that was able to output mostly correct results but also some that, while plausible, were incorrect.
(3) When a participant selected a video span, it is possible that either the start or end of that span would not perfectly overlap the annotated timeframe of a movement in the video. To address this, the prototype classified a movement as being within a given span only if more than half of its duration fell within the span's boundaries.
(4) Since a selected span could include multiple dance movements, the list of dictionary-search results displayed combined results from the match lists for all movements within that span, as follows: From among the match lists for all movements in the span, we first presented the matches for all the moves, in a randomized order. We then presented the closest matches for all the movements, also in randomized order. This final list was capped at 50 moves.

5.3 Baseline Design

The baseline prototype (part of the same web application) had exactly the same design as the search prototype except there was no search panel and "search in selection" button. We also provided access to an external feature-based Cuban Salsa dictionary system to our participants [39]. Figure 6 described our tools interface for the baseline condition and the feature-based dictionary.

5.4 Experimental Study Participants

We posted an advertisement across various university dance groups and shared it on the Salsa dance subreddit (r/Salsa). The advertisement included two questions: "Are you currently involved in Salsa dance or choreography?" and "Have you been engaged in Salsa dance or choreography in the past 5 years?" We invited individuals who answered yes to at least one question to participate in the study. In total, we recruited eight dance learners, 6 women, 1 man, and 1 non-binary individual. The median age of participants was 25 (σ=8.07). Five participants reported having more than 3 years of experience in dance, one reported between 1 and 2 years, and 2 reported less than a year. Two of our participants had more than 3 years of experience in choreographing or performing Cuban Salsa, three had between 1 and 2 years of experience, and two had less than a year of experience.

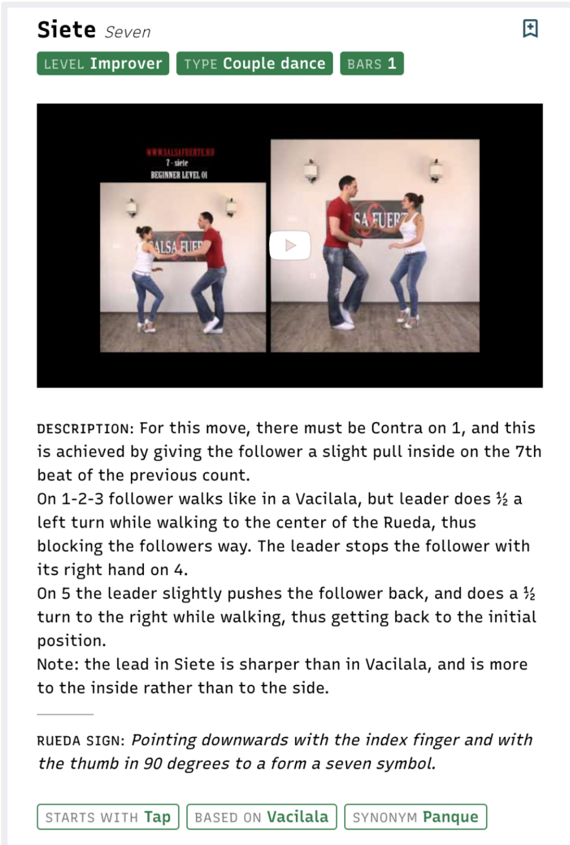


Figure 5: Screenshot of an individual result snippet. It displays the name of the dance move in bold at the top along with the English translation. At the top-right corner, a bookmark-icon-shaped button allowed participants to automatically add an annotation to the position of the current span selection with information about this particular dance move. Underneath there is information on the level, Type, and Bars (representing difficulty) of the move. There is a video of the dance move with performance followed by a description following the commonly used nine-step breakdown. At the bottom, there is the Rueda sign (symbolic descriptions of moved), what the movement starts with, what it is based on, and a synonym (closely related move).

5.5 Experimental Study Procedure

At the beginning of this online study, participants were presented with an IRB-approved informed consent form to review and sign before proceeding. Following this, we distributed a demographics and past experiences form to the participants for completion. Each participant was assigned a unique two-digit ID (P11-P18), which was used to determine the sequence of videos they would watch and script for. A total of six videos were viewed by each participant, with three videos featuring an in-situ auto-identify functionality and the other three without it. The order of the videos was modified for each participant. The first sub-group of four participants watched and wrote notes for the three videos without the auto-identify feature, while the second sub-group of four participants watched and wrote notes for the three videos with the auto-identify feature. In our instructions, we asked the participants to draft notes enabling them to execute the sequence of dance moves independently at a later time, given they are provided with just the notes and glossary of Cuban Salsa dance moves. After writing this script for each video, participants were asked to complete a NASA TLX questionnaire [45, 46], and provide open-ended feedback. Their interactions with the prototype system were recorded:

(1) The notes and individual annotations on spans, as well as any modifications made to them, were automatically recorded.
(2) Our prototype recorded the starting and ending points on the video timeline of every span the participant selected, the number of signs within each span, whenever the participant triggers a search, the amount of time spent watching each video, and the text of the description typed by the participant.
(3) A researcher took observational notes during the experiment based on the interactions observed on the participants' shared screens.

An analysis of the interaction data listed above was performed by a member of our research team who reviewed and coded this data from the perspective of identifying typical sequences of interaction behavior during each video session. They reviewed recordings of the screen, analyzed data captured by the software prototype, and reviewed the observer's notes. At the end of the entire experiment, a debriefing interview was conducted to gather participants' impressions of the system, perceptions of how they interacted with the device, and other recommendations. The interview data was transcribed and coded using a deductive approach based on the interaction categories identified.

6 EXPERIMENTAL STUDY FINDINGS

6.1 Notes Quality

To assess the quality of the notes, an expert-certified Cuban Salsa choreographer and instructor read the descriptions and assigned overall scores (out of 10). They first watched all the videos and wrote notes to use as a reference later. The order of the notes was randomized, and the judge did not know which notes were produced using the auto-identify prototype and which had been produced using the baseline prototype. The judge evaluated the dance notes provided in a spreadsheet, primarily focusing on the accuracy of named moves. This involved scrutinizing each step for precision, noting omissions and additions, and subjectively evaluating the quality of textual descriptions. The textual descriptions covered transitions, non-standard move names, and instances where participants resorted to text when they could not find the name of the desired move. After tabulating these elements, a holistic score was assigned. The average note quality was 8.71 (σ=2.45) for the auto-identify prototype and 5.65 (σ=3.13) for the baseline. The distributions in scores between the two conditions varied significantly (Mann-Whitney U = 10, n_{Baseline} = 8, n_{Salsa} = 8, p = 0.032 < 0.05, two-tailed, moderate effect).

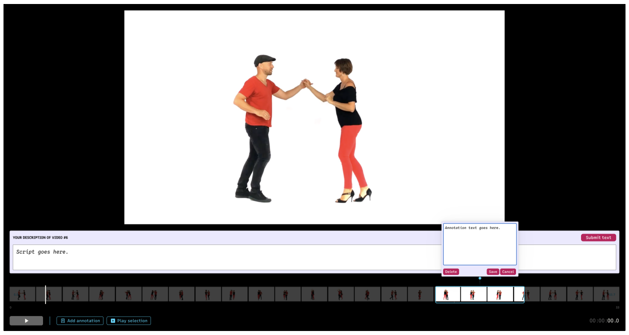
(a)  (b) 

Figure 6: (a) The prototype utilized in the baseline condition of Study 2 is identical to the Search Prototype, with the exception that it lacked any search capability. (b) The Salsayo Salsa move search website, which participants used in the Study 2 baseline condition. As users click on different properties, a list of dance moves appears that can be clicked to watch.

TLX Sub-Scale	Search Prototype	Baseline Prototype	Significance Testing
Mental Demand	51.5	58.0	p=0.757, U=36.5
Physical Demand	33.3	39.8	p=0.596, U=34
Temporal Demand	46.9	50.5	p=0.441, U=35
Effort	47.6	52.5	p=0.265, U=37
Performance	29.3	60.6	p=0.00714, U=9.5 **
Frustration	37.2	56.0	p=0.0021, U=10 **

Table 2: NASA TLX sub-scale scores from participants in the dictionary-search and baseline prototype conditions, with scores scaled to a 0-to-100 range. For all sub-scales, lower scores are better, i.e., indicating less perceived demand, less effort needed, less frustration, or a better sense of performance success. The significance testing column displays results from two-tailed Mann-Whitney U tests. (indicates $p < 0.05$)**

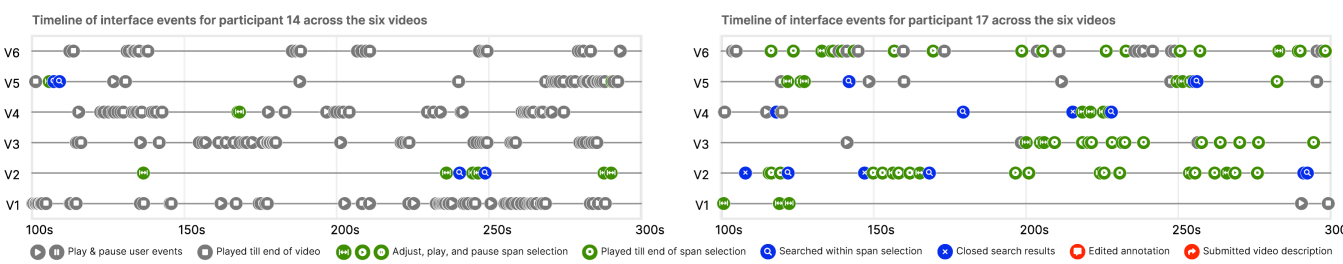


Figure 7: This chart visually represents a chronological subset of the sequence of events undertaken by participants 14 and 17 while working on each video. Although the video order was randomized for each participant, we have presented it in sequential order here to facilitate comparisons. Each action is associated with a specific icon, as indicated in the legend. However, due to the occasional high density of actions resulting in impenetrable overlays, we have also grouped actions by color to help identify overall patterns. The auto-identify feature was enabled for Videos 2, 4, and 5.

6.2 Workload

We compared the workload associated with note-taking across both tasks since one of the key motivations of our work was to streamline the process of looking up unfamiliar dance moves for learners, something that was found to be time-consuming and generally lacking for current approaches.

Table 2 contains mean scaled NASA-TLX sub-score values (physical demand, temporal demand, performance, effort, and frustration) s[45] and results of two-tailed Mann-Whitney U tests comparing sub-scores across two conditions. Participants' responses when using the auto-identify prototype had significantly lower values[1] for frustration (how insecure, discouraged, irritated, or stressed they felt) and performance (how successful they were in what they were asked to do).

6.3 Analysis of Observations

6.3.1 Using Span-selector to Constrain the Playhead. Seven out of eight participants used the span-searching tool to constrain the portion of the video that would play at one time. The utilization of

[1]For all sub-scales, lower values indicate better outcomes.

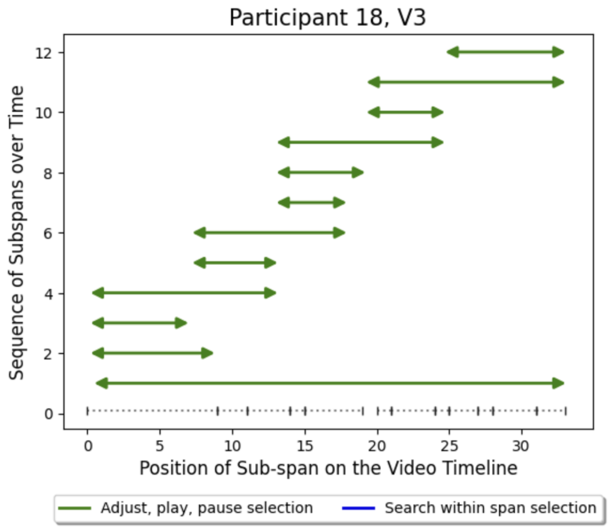


Figure 8: P18 viewed a video in a linear manner, using the span selection to progressively view short video segments. Horizontal bars indicate spans selected with respect to the total duration of the video. The first span selected is at the bottom of the y-axis, with subsequent spans appearing higher on the y-axis. The black lines at the bottom show the actual signs on the video timeline.

span selection varied among different participants, as demonstrated in Figure 7, where we highlight two distinct usage patterns. Both P14 and P17 were aware of and actively used span selection. However, while P14 primarily employed it as a tool to limit the search input, P17 not only utilized it for this purpose but also extensively employed span selection to constrain the playhead's movements.

Figure 8 also illustrates this trend for P18's interactions with a video while using the baseline prototype. It plots the positions of spans selected over time as they viewed different portions of a video incrementally. Span 1, shown at the bottom of the figure, indicates the first span selected. The horizontal line with dividers underneath shows the actual location of the dance moves on the video timeline. We can see from the chart that P18 engaged in a *fine-tuning* behavior with the duration of the spans.

The median duration of spans for participants who only used the playhead to watch the video for the auto-identify condition, was 13.21 seconds (σ = 15.79 seconds), equal to 2.986 dance moves.

6.3.2 Using the auto-identify feature to inform notes. In 18 out of 24 video sessions (75%) where participants had access to the auto-identify feature, they made use of the auto-identify feature to look up dance moves in the video. As illustrated in Figure 9, the results of the auto-identify tool informed participants' translation decisions as they progressed through the video.

Participants also commented favorably about the contents of the results. P11 liked the information in the result snippet: *"It gave me ample resources to explain a basic partner Salsa move."* Similarly, P13 said: *"I think that the search bar helped the most for this video because I was so unfamiliar with some moves, and they were very similar, which made it difficult, but seeing the moves separately made it easier to put together."* P12 also mentioned liking the search feature: *"I like the idea of a tool that enables dancers to search for unfamiliar moves to assist in their ability to understand videos – whether social dance clips or demos from a dance studio or workshop – and learn new skills."*

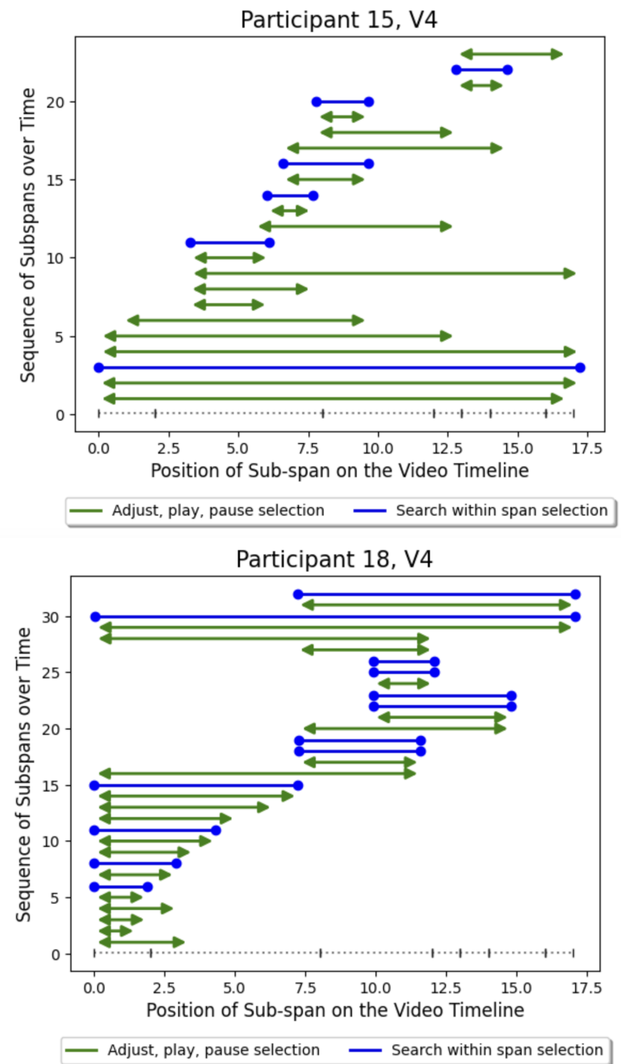


Figure 9: (a) Blue lines indicate spans for which search was requested, and green solid lines indicate spans for which no search was requested. P15 watching the entire video, watching some segments, and then performing searches on some of the spans. (b) P18 gradually increases and decreases the size of the span to re-watch segments and make a search.

Participants, including P11, particularly liked the step-wise detailed description of the steps that were provided in the result snippets: *"The search function definitely assists in using the correct terminology and maybe providing clearer analogies to describe the movement rather than just describing the directionality. By having the* *links to sample videos that break it down, the movement will become easier to learn."*

Additionally, two participants also mentioned that they used the search feature to confirm sections that did not have a named dance move before proceeding to use the annotation feature to annotate those segments. P17 commented on this: *"It also helps learners to decipher which parts are just freestyle (ex. the little hop at the end) as the search doesn't specify move names for it."*

6.3.3 Adjusting the span durations before making searches. We observed participants adjusting the durations of spans both upwards and downwards before conducting searches. In Figure 9, you can see examples of participants modifying the span length. Specifically, at least three participants expanded the size of their spans before commencing a search. Additionally, some participants exhibited a *fine-tuning* behavior, making numerous adjustments before finalizing the span for the auto-identify function. Figure 7 illustrates how both P14 and P17 fine-tuned their span boundaries prior to initiating a search, as evident in their actions with video 2.

We also analyzed how close the selected spans were when the search was made to the actual positions of moves in the videos. The median error between each span selection boundary and the nearest actual sign boundary was 1.25 seconds. For reference, the median duration of a dance move in our study was 3.27 seconds.

6.3.4 Challenges with disambiguation between results. We observed that at least four of our participants struggled with disambiguating between results at least once during the experiment. In the open-ended feedback and debriefing interview, three of them mentioned that it was challenging to disambiguate between similar results. For example, P18 stated: *"[It was] easy to search, but it was difficult to distinguish between some of the moves since so many of them overlap."* P13 reported that they read the descriptions first and then watched the videos, and that helped them browse the results list quicker if the correct result was not at the top: *"It was good [that] the [results] videos had a description, and when I clicked the search, it matched."* Two participants, including P15, also mentioned that it was difficult to match the span-segment in the video with the video in the result snippet: *"It was definitely helpful, but the moves [results] were often a bit different."*

Some of the participants also thought that 50 results are too many: *"It was not hard to select sub-spans of the video to rewatch, but again, the resulting hits numbered 50, which may be too time-consuming to weed through."* One participant (P14) commented that they would only like to see the top result: *"I want to click on the move and have the search tool pop up and tell me what that particular move was at that moment and then give me the breakdown by counts then with music so I can understand what I just saw."*

7 DISCUSSION

Although prior research indicates that dance learners frequently utilize videos in their learning process [42, 51, 102], little attention has been given to understanding the video-watching and note-taking habits of dance learners. Our interview study honed in on the engagements of dance learners with videos, shedding light on their video-watching and note-taking practices. We explored how users interact with video players, particularly when navigating through difficult dance sequences, and identified the workaround strategies they employ to facilitate comprehension of the dance video, such as periodic pausing, reducing playback speed, and rewinding. Furthermore, our study underscores the importance placed on note-taking for dance learners by revealing patterns in note-taking and structuring that involved breaking down dance performance sequences into segments, individual moves, and steps. Finally, we found a significant interest among certain viewers in learning the formal names of dance moves and seeking out additional relevant information, particularly in the context of formal dance styles.

The interview study yielded two important insights: firstly, the pressing necessity to facilitate novice dance learners in compiling detailed and informative notes, and secondly, assisting them in identifying formal dance move names without disrupting their video viewing and comprehension experience. As highlighted in the related work section, choreographers' notebooks serve as vital resources in aiding learners during the note-taking process [2, 15, 18, 19, 32, 35, 101, 112]. Despite their benefits, these tools do not support users in looking up unfamiliar dance moves while taking notes on videos. Our findings revealed current methodologies aiding users in identifying unfamiliar dance moves predominantly involve utilizing tags and feature-based catalogs, among other strategies. However, users still face challenges while describing intricate dance moves textually and reconciling different naming conventions across different regions. This motivated an in-situ tool that uses video-span as an input to search, devised to empower novice dance learners to enhance their note-taking and dance move auto-identification capabilities without deterring their engagement with the video content.

Our second study was a comparative experimental study between our note-taking and span-based dance move auto-identification tool and a baseline prototype. The latter also offered span selection functionality to restrict the playhead, yet required dance learners to use an external, pre-existing tool for referencing dance moves [39]. Findings demonstrated that the use of an integrated auto-identify tool led to higher-quality notes and a decrease in perceived workload. Notably, two out of the five metrics on the NASA TLX (namely Frustration and Performance) were significantly different across both conditions. An exploration of behavioral data unveiled distinct usage patterns. Notably, we found that the playhead was used to constrain the video timeline and modify it to determine a suitable unit of analysis. Additionally, the auto-identify feature was utilized to inform notes and identify segments without formal dance moves. However, disambiguating between long videos of results remains a challenge for some participants.

7.1 Supporting Users during Different Stages of Dance Learning

A previous interview study with professional dance learners conducted a task analysis of the dance learning process [93]. This study pinpointed three fundamental operations—analysis, integration, and personalization—that were consistently observed across participants. Notably, all participants identified the initial step in the preparation phase as observation, which often involves scrutinizing isolated components. While the participants in the first study did not express a preference for employing formal nomenclature during this process, they did emphasize a focus on discerning and assimilating individual dance elements.

An essential aspect of the analysis phase often involves segmentation operation [31, 60, 78], simplifying complex movements to facilitate easier understanding and imitation before introducing creative modifications. In dance research, complexity is characterized by the granularity of details discerned within a movement, escalating with a more generalized or coarse analysis [80, 93]. The utilization of a span-based playhead allows users to progressively modify their unit of analysis, transitioning from scrutinizing small sequences to examining individual dance moves and, ultimately, the steps encompassed within a specific move. Interestingly, in the second study, we observed two contradictory user behaviors concerning span adjustment: subsequent decrease in the spans they analyzed as expected, but also a gradual increase until a search is made. These patterns underscore that users frequently need to review varying lengths of video to accurately gauge the length of dance video that they would like to analyze at one time. Moreover, since dance moves often fluidly merge into one another [7], users tend to use auto-identify feature for pauses or demarcations between recognizable segments as a precursor to a detailed analysis.

Observation in dance is an ongoing, iterative process throughout the learning journey. The benefits of looking-up an unfamiliar dance move in a video extend beyond initial instruction. If a dancer faces challenges with a specific segment later on, revisiting the video allows for close examination and improved understanding, enhancing practice.

Knowing formal dance names aids effective practice by allowing learners to focus on perfecting one named move at a time. These names are like the alphabet, serving as "building blocks" for learning, creating, or modifying choreographies. Tracking progress becomes simpler; learners can set goals, noting mastered moves and those still in progress. Knowledge of names facilitates pinpointing areas for correction and fine-tuning, aiding focused improvements. Although our study did not explore our tool's potential for personalization and creative expression, it suggests that familiarizing learners with diverse dance moves could enhance creativity. This tool goes beyond offering sequences, promoting exploration and learning among dance learners.

7.2 Informing Design, Computer Vision, and Dance Researchers

Our findings also provide **design recommendations** for dance video comprehension tools that feature a video-span based user-interface:

(1) The concept of timeline-based video annotation has been previously proposed in the context of annotation frameworks for generating reliable ground-truth datasets for dance [37]. Here, we offer insights into its implementation in a video comprehension tool specifically designed for learners. We suggest integrating a timeline-based video playhead that showcases sequential captures extracted directly from the video. While not intended to replace the conventional linear video timeline functionality, allowing users to traverse the video from start to end, our observations reveal distinct benefits for various user groups with these two timeline navigation approaches. Unlike a standard traditional timeline, the playhead empowers users to examine specific segments of the entire video, iterating over different self-constrained sections for a detailed exploration. Despite the rising popularity of this UI paradigm in video editing systems, exemplified by YouTube's clip feature, its application in the context of dance video comprehension tools remains largely unexplored [50].

(2) Prior work has suggested incorporating features for attaching notes to a "point-based" timeline and a separate journaling feature [101]. We propose incorporating two note-taking features within the timeline-based video playhead: one that attaches a note directly to the timeline and another for general note-taking. Additionally, we recommend a mechanism to automatically transfer metadata from the auto-identified dance move into the timeline annotation.

(3) Comprehensive meta-data encompassing aspects such as difficulty level, engaged body parts, similar dance moves, and the number of dancers involved should be readily accessible to users in the result snippet. Inspired by prior work on using common verbs in computational choreography systems [53] and guided by our interview study, we included brief descriptions for each dance move. These were favorably received by participants, and as such we recommend future designers adopt this approach.

(4) We recommended adding textual descriptions of the dance moves that use standardized numbering or other established conventions to systematically break down each move into individual steps.

(5) Whenever possible, videos showcased in the results should offer footage captured from various angles.

For dance move recognition researchers, our research introduces a novel task: utilizing a segment of a continuous dance video as a search input, in contrast to the traditional approach of using a complete video. This approach holds the potential to augment the accuracy of recognition results by capitalizing on users' perceptions of dance movement boundaries, which recent research has shown to be of impressive quality [78]. Our observational findings from the second study indicated that the median error for span selection boundaries was 1.25 seconds, which is, on average, less than 38% of an average Cuban Salsa dance move.

Importantly, this approach also encourages learners to independently analyze challenging dance videos, since they are aided by on-demand in-situ sign-searching instead of an automatically generated description of the video taken as a whole. However, as identified in other similar human-movement contexts [1, 48, 64], such recognition tasks still face challenges: the extracted fragment might not align with movement boundaries, may exhibit co-articulation effects from dance moves outside the selected span, or offer little contextual information for the recognition system.

Findings from the second study also uncovered the potential use of video span as a useful probe to understand how dance learners segment videos, potentially giving insights about how learners are breaking down complex videos [5, 9]. Our observational findings revealed how these span selections were frequently fine-tuned by our participants, to conduct better searches or to constrain the video playhead along dance move boundaries. Dance researchers could use this UI paradigm to investigate nuanced concepts in dance, e.g., coarticulation [7].

7.3 Implications of our Findings for Other Forms of Dance and Performance Contexts

The interpretation of dance poses varies across different styles, and is intertwined with cultural and societal narratives [100, 111]. The unique characteristics inherent in various dance styles necessitate the development of comprehension tools specifically designed to accommodate the distinct nuances of each style. While certain computer vision approaches could excel at recognizing overarching dance styles [82], systems geared towards identifying specific moves for glossary-based styles typically demand individually-tailored models [56].

More broadly, our research speaks to the literature on frameworks and tools that support the annotation and analysis of video sequences of human performance, such as aerobic exercises and sports[108, 118], surgery [40, 40, 75, 87, 117], music [77], and cooking [8], language transcription, including signed languages [11, 47–49, 73], etc. These frameworks are frequently used by annotators for data collection to aid in human action recognition initiatives. In addition, videos are also analyzed and annotated in other contexts, either to support learning (e.g., junior surgeons watching surgery videos for educational purposes [34]) or to critically evaluate human performances in videos (e.g., judges assessing gymnasts to score their performances [118]).

However, the findings of prior work, suggest an absence of movement look-up tools integrated within the frameworks. Thus, users viewing various types of human-movement videos to identify components of human movements often employ some of the workarounds identified in our study. Our findings contribute to the broader literature on searching through video spans of human performance, which currently heavily relies on external feature-based catalogs. These findings may be particularly relevant in contexts where structured sequences of named human movements are involved and could motivate the design of span-based in-situ video look-up tools.

8 LIMITATIONS AND FUTURE WORK

This work has several limitations of this work that can serve as avenues for future research:

(1) In Study 1, the majority of our participants were at an early stage in their dance learning journey and were younger than 22 years old. We chose this participant pool because video analysis is something that is encouraged to early dance learners and they are likely to find more challenges with memorization. Future studies could also recruit participants from a more diverse age range and expert dancers, choreographers, and studio heads who specialize in different dance styles to gain insights into the note-taking practices.

(2) In our prototype design, our focus was on incorporating the search functionality in the best possible way. There are several other features commonly available in choreographers' notebooks mentioned in the related work section, e.g., note-taking on the video itself, motion analysis, audio descriptions, etc. Future researchers could investigate how to add those features to a video comprehension tool with in-situ dance move search functionality.

(3) Our choice of focusing on Cuban Salsa in Study 2 was informed by the participants' interest in note-taking for Salsa and the existence of a web-based, feature-based dictionary that could be used as a baseline. Future researchers can adapt our tool for other dance styles where catalogs of dance moves exist, e.g., Ballet [109].

(4) Although Study 2 revealed benefits in script quality for participants, due to time constraints of an experimental study, we could not examine whether the tool's use led to a better kinesthetic learning and subsequent performance. Future research could investigate the short-term and long-term benefits of using this technology, with a focus on improving learning and performance.

(5) Extensive prior work described in subsection 2.3.1 has focused on computational choreography. However, we did not explore adding a feature to our system where the system automatically composes a choreographic sequence based on the notes and labels provided by the user. Future researchers could explore the benefits of this technology. Potentially, a composition tool in this context could aid in the verification of labels and allow users to cross-compare the actual video with a video composed based on their annotations.

(6) Although not highlighted in the findings, several participants mentioned how the tool could be transformed into a social application where dancers can receive feedback from their peers and share resources. Future research can investigate the design of social technologies to support the identification of dance moves.

9 CONCLUSION

Automatic segmentation and recognition of a sequence of moves in a continuous dance video remain a challenging task. However, given recent advancements in computer vision, it might be possible in the near future for a system to provide the closest matching dance moves to a segment of video submitted by users.

In this paper, we explore the design of a technology that enables learners to select a span of a video that is then used as an input query into a dance-move-identification system, which outputs dance moves within it as they engage in video analysis and note-taking tasks. Findings from our interview study underscored the importance of note-taking in this process and participants' interest in learning the formal names and other details about the dance moves. Leveraging these findings, we developed a tool to assist dance learners in video comprehension of Cuban Salsa. In a subsequent experimental study, we evaluated this tool, and uncovered benefits including enhanced note quality and reduced workload. Our observational findings also provided insights into users' utilization of the span-selector and auto-identify features.

For developers creating interactive tools for dance, our findings provide evidence supporting the benefits of integrating a span-selector to constrain and analyze videos, offer insights into potential usage patterns, and suggest design recommendations. Additionally, for computer vision researchers, these results present a novel task involving user-selected spans as inputs for the auto-identification of dance moves in forms adhering to a predefined glossary of moves.

ACKNOWLEDGMENTS

This material is based upon work supported by the Department of Health and Human Services under Award 90DPCP0002-0100; by the National Science Foundation under Grant No. 1954284, 2125362, 2212303, and 2235405; and by the Fulbright Commission (Fulbright-CAPES Scholarship, ME / CAPES N° 8 / 2020). We would like to thank the Latin Rhythm Dance Club at Rochester Institute of Technology and Faith Avilene DePano for feedback on different parts of this project.

The authors acknowledge the use of OpenAI's ChatGPT and Google's Bard as editorial tools to refine the clarity and style of this manuscript's presentation.

REFERENCES

[1] Adam Agassi, Hadas Erel, Iddo Yehoshua Wald, and Oren Zuckerman. 2019. Scratch Nodes ML: A Playful System for Children to Create Gesture Recognition Classifiers. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3290607.3312894

[2] Omid Alemi, Philippe Pasquier, and Chris Shaw. 2014. Mova: Interactive Movement Analytics Platform. In *Proceedings of the 2014 International Workshop on Movement and Computing* (Paris, France) *(MOCO '14)*. Association for Computing Machinery, New York, NY, USA, 37–42. https://doi.org/10.1145/2617995.2618002

[3] Äli Leijen, Ineke Lam, Liesbeth Wildschut, P. Robert-Jan Simons, and Wilfried Admiraal. 2009. Streaming video to enhance students' reflection in dance education. *Computers & Education* 52, 1 (Jan. 2009), 169–176. https://doi.org/10.1016/j.compedu.2008.07.010

[4] Maria João Alves. 2017. The use of video as self-evaluation in dance classes. In *Exploring Identities in Dance: Proceedings from the 13th World Congress of Dance and the Child International*. University of Copenhagen Denmark, University of Libson, Portugal, 5–10.

[5] Egil Bakka and Gediminas Karoblis. 2010. Writing A Dance: Epistemology for Dance Research. *Yearbook for Traditional Music* 42 (2010), 167–193. https://doi.org/10.1017/S0740155800012704

[6] Liesbeth Pankaja Bennink, Kandhan Raja Deekshithar, Jayakumar Raja Deekshithar, and Sankar Raja Deekshithar. 2013. Shiva's Karanas in the temples of Tamil Nadu: the Natya Shastra in stone.

[7] Frédéric Bevilacqua, Baptiste Caramiaux, and Jules Françoise. 2016. Perspectives on Real-time Computation of Movement Coarticulation. In *Proceedings of the 3rd International Symposium on Movement and Computing* (Thessaloniki, GA, Greece) *(MOCO '16)*. Association for Computing Machinery, New York, NY, USA, Article 35, 5 pages. https://doi.org/10.1145/2948910.2948956

[8] Simone Bianco, Gianluigi Ciocca, Paolo Napoletano, Raimondo Schettini, Roberto Margherita, Gianluca Marini, Giorgio Gianforme, and Giuseppe Pantaleo. 2013. A semi-automatic annotation tool for cooking video. In *Image Processing: Machine Vision Applications VI*, Vol. 8661. SPIE, Bicocca Open Archive, University of Milano-Bicocca, Bicocca, Italy, 317–326.

[9] Johannes Birringer. 2004. Dance and Interactivity. *Dance Research Journal* 36, 1 (2004), 88–111. https://doi.org/10.1017/S0149767700007580

[10] Julien Blanchet, Megan Hillis, Yeongji Lee, Qijia Shao, Xia Zhou, David Kraemer, and Devin Balkcom. 2023. Automatic Generation and Teaching of Dance Lessons from Video. In *Proceedings of the 24th International Workshop on Mobile Computing Systems and Applications* (Newport Beach, California) *(HotMobile '23)*. Association for Computing Machinery, New York, NY, USA, 145. https://doi.org/10.1145/3572864.3581592

[11] Matyas Bohacek and Saad Hassan. 2023. Sign Spotter: Design and Initial Evaluation of an Automatic Video-Based American Sign Language Dictionary System. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (New York, NY, USA) *(ASSETS '23)*. Association for Computing Machinery, New York, NY, USA, Article 92, 5 pages. https://doi.org/10.1145/3597638.3614497

[12] James H Bradford and Paulette Côté-Laurence. 1995. An application of artificial intelligence to the choreography of dance. *Computers and the Humanities* 29 (1995), 233–240.

[13] David Byrne. 2022. A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Quality & quantity* 56, 3 (2022), 1391–1412.

[14] Diogo Cabral, Urândia Carvalho, João Silva, João Valente, Carla Fernandes, and Nuno Correia. 2011. Multimodal Video Annotation for Contemporary Dance Creation. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI EA '11)*. Association for Computing Machinery, New York, NY, USA, 2293–2298. https://doi.org/10.1145/1979742.1979930

[15] Diogo Cabral, João G. Valente, Urândia Aragão, Carla Fernandes, and Nuno Correia. 2012. Evaluation of a multimodal video annotator for contemporary dance. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (Capri Island, Italy) *(AVI '12)*. Association for Computing Machinery, New York, NY, USA, 572–579. https://doi.org/10.1145/2254556.2254663

[16] Kelly Caine. 2016. Local Standards for Sample Size at CHI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 981–992. https://doi.org/10.1145/2858036.2858498

[17] Tom W. Calvert, Armin Bruderlin, Sang Mah, Thecla Schiphorst, and Chris Welman. 1993. The Evolution of an Interface for Choreographers. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems* (Amsterdam, The Netherlands) *(CHI '93)*. Association for Computing Machinery, New York, NY, USA, 115–122. https://doi.org/10.1145/169059.169113

[18] Antonio Camurri, Shuji Hashimoto, Matteo Ricchetti, Andrea Ricci, Kenji Suzuki, Riccardo Trocca, and Gualtiero Volpe. 2000. Eyesweb: Toward gesture and affect recognition in interactive dance and music systems. *Computer Music Journal* 24, 1 (2000), 57–69.

[19] Kristin Carlson, Herbert H. Tsang, Jordon Phillips, Thecla Schiphorst, and Tom Calvert. 2015. Sketching Movement: Designing Creativity Tools for in-Situ, Whole-Body Authorship. In *Proceedings of the 2nd International Workshop on Movement and Computing* (Vancouver, British Columbia, Canada) *(MOCO '15)*. Association for Computing Machinery, New York, NY, USA, 68–75. https://doi.org/10.1145/2790994.2791007

[20] Daniel Castro, Steven Hickson, Patsorn Sangkloy, Bhavishya Mittal, Sean Dai, James Hays, and Irfan Essa. 2018. Let's dance: Learning from online dance videos.

[21] Caroline Chan, Shiry Ginosar, Tinghui Zhou, and Alexei Efros. 2019. Everybody Dance Now. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE, IEEE Corporate Headquarters (NY). 3 Park Avenue, 17th Floor, 5932–5941. https://doi.org/10.1109/ICCV.2019.00603

[22] Jacky CP Chan, Howard Leung, Jeff KT Tang, and Taku Komura. 2010. A virtual reality dance training system using motion capture technology. *IEEE transactions on learning technologies* 4, 2 (2010), 187–195.

[23] Diane Chi, Monica Costa, Liwei Zhao, and Norman Badler. 2000. The EMOTE model for effort and shape. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '00)*. ACM Press/Addison-Wesley Publishing Co., USA, 173–182. https://doi.org/10.1145/344779.352172

[24] Leigh Clark, Philip Doyle, Diego Garaialde, Emer Gilmartin, Stephan Schlögl, Jens Edlund, Matthew Aylett, João Cabral, Cosmin Munteanu, Justin Edwards, et al. 2019. The state of speech in HCI: Trends, themes and challenges. *Interacting with computers* 31, 4 (2019), 349–371.

[25] Roger Copeland and Marshall Cohen. 1983. *What is dance?: Readings in theory and criticism*. Vol. 720. Oxford [Oxfordshire]; New York: Oxford University Press, Kettering, Northamptonshire. NN14 1UA.

[26] Michael Coyle and Tom Calvert. 2002. Choreography and Animation: Tools for Visualization of Dance and Notation. , 6 pages.

[27] Changrui Cui, Junhuai Li, Dandan Du, Huaijun Wang, Pengjia Tu, and Ting Cao. 2021. The Method of Dance Movement Segmentation and Labanotation Generation Based on Rhythm. *IEEE Access* 9 (2021), 31213–31224. https://doi.org/10.1109/ACCESS.2021.3060103

[28] Aspasia Dania, Dimitrios Hatziharistos, Maria Koutsouba, and Vasiliki Tyrovola. 2011. The use of technology in movement and dance education: Recent practices and future perspectives. *Procedia-Social and Behavioral Sciences* 15 (2011), 3355–3361.

[29] Mampi Devi, Sarat Saharia, and D.k.bhattacharyya. 2015. Article: Dance Gesture Recognition: A Survey. *International Journal of Computer Applications* 122, 5 (July 2015), 19–26.

[30] Ragini Devi. 1990. *Dance dialects of India*. Motilal Banarsidass Publ., Naraina Industrial Area, Phase-1, New Delhi, India.

[31] Paula M Di Nota, Michael P Olshansky, and Joseph FX DeSouza. 2020. Expert event segmentation of dance is genre-specific and primes verbal memory. *Vision* 4, 3 (2020), 35.

[32] Augusto Dias Pereira dos Santos, Lian Loke, and Roberto Martinez-Maldonado. 2018. Exploring Video Annotation as a Tool to Support Dance Teaching. In *Proceedings of the 30th Australian Conference on Computer-Human Interaction* (Melbourne, Australia) *(OzCHI '18)*. Association for Computing Machinery, New York, NY, USA, 448–452. https://doi.org/10.1145/3292147.3292194

[33] Sally Doughty, Kerry Francksen, Michael Huxley, and Martin Leach. 2008. Technological enhancements in the teaching and learning of reflective and creative practice in dance. *Research in Dance Education* 9, 2 (2008), 129–146.

[34] Shahram Eivazi, Roman Bednarik, Markku Tukiainen, Mikael von und zu Fraunberg, Ville Leinonen, and Juha E. Jääskeläinen. 2012. Gaze Behaviour of Expert

and Novice Microneurosurgeons Differs during Observations of Tumor Removal Recordings. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (Santa Barbara, California) *(ETRA '12)*. Association for Computing Machinery, New York, NY, USA, 377–380. https://doi.org/10.1145/2168556.2168641
[35] Katerina El Raheb, Michele Buccoli, Massimiliano Zanoni, Akrivi Katifori, Aristotelis Kasomoulis, Augusto Sarti, and Yannis Ioannidis. 2023. Towards a general framework for the annotation of dance motion sequences: A framework and toolkit for collecting movement descriptions as ground-truth datasets. *Multimedia Tools and Applications* 82, 3 (2023), 3363–3395.
[36] Katerina El Raheb, Aristotelis Kasomoulis, Akrivi Katifori, Marianna Rezkalla, and Yannis Ioannidis. 2018. A Web-Based System for Annotation of Dance Multimodal Recordings by Dance Practitioners and Experts. In *Proceedings of the 5th International Conference on Movement and Computing* (Genoa, Italy) *(MOCO '18)*. Association for Computing Machinery, New York, NY, USA, Article 8, 8 pages. https://doi.org/10.1145/3212721.3212722
[37] Katerina El Raheb, Nicolas Papapetrou, Vivi Katifori, and Yannis Ioannidis. 2016. BalOnSe: Ballet Ontology for Annotating and Searching Video Performances. In *Proceedings of the 3rd International Symposium on Movement and Computing* (Thessaloniki, GA, Greece) *(MOCO '16)*. Association for Computing Machinery, New York, NY, USA, Article 5, 8 pages. https://doi.org/10.1145/2948910.2948926
[38] Angel Fiallos, Carlos Fiallos, and Stalin Figueroa. 2021. Tiktok and Education: Discovering Knowledge through Learning Videos. In *2021 Eighth International Conference on eDemocracy & eGovernment (ICEDEG)*. IEEE, Piscataway, New Jersey, 172–176. https://doi.org/10.1109/ICEDEG52154.2021.9530988
[39] Ido Flaishon. 2022. Salsa Yo - Watch & Dance. https://salsayo.com/
[40] Carly R Garrow, Karl-Friedrich Kowalewski, Linhong Li, Martin Wagner, Mona W Schmidt, Sandy Engelhardt, Daniel A Hashimoto, Hannes G Kenngott, Sebastian Bodenstedt, Stefanie Speidel, et al. 2021. Machine learning for surgical phase recognition: a systematic review. *Annals of surgery* 273, 4 (2021), 684–693.
[41] Abraham Glasser, Matthew Watkins, Kira Hart, Sooyeon Lee, and Matt Huenerfauth. 2022. Analyzing Deaf and Hard-of-Hearing Users' Behavior, Usage, and Interaction with a Personal Assistant Device That Understands Sign-Language Input. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 306, 12 pages. https://doi.org/10.1145/3491102.3501987
[42] Dimitra Gratsiouni, Maria Koutsouba, Foteini Venetsanou, Vasiliki Tyrovola, et al. 2016. Learning and digital environment of dance–the case of Greek traditional dance in YouTube. *European Journal of Open, Distance and E-learning* 19, 2 (2016), 80–95.
[43] Carla F. Griggio, Mario Romero, and Germán Leiva. 2015. Towards an Interactive Dance Visualization for Inspiring Coordination Between Dancers. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI EA '15)*. Association for Computing Machinery, New York, NY, USA, 1513–1518. https://doi.org/10.1145/2702613.2732925
[44] Kelly Hamilton. 2015. Mapping motion II: motion capture and the visualisation of dance. In *Proceedings of the Conference on Electronic Visualisation and the Arts* (London, United Kingdom) *(EVA '15)*. BCS Learning & Development Ltd., Swindon, GBR, 46–55. https://doi.org/10.14236/ewic/eva2015.5
[45] Sandra G Hart. 2006. NASA-task Load Index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage Publications Sage CA, Sage publications, Los Angeles, CA, 904–908.
[46] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, Amsterdam, Netherlands, 139–183.
[47] Saad Hassan, Oliver Alonzo, Abraham Glasser, and Matt Huenerfauth. 2021. Effect of Sign-recognition Performance on the Usability of Sign-language Dictionary Search. *ACM Trans. Access. Comput.* 14, 4, Article 18 (oct 2021), 33 pages. https://doi.org/10.1145/3470650
[48] Saad Hassan, Akhter Al Amin, Caluã de Lacerda Pataca, Diego Navarro, Alexis Gordon, Sooyeon Lee, and Matt Huenerfauth. 2022. Support in the Moment: Benefits and Use of Video-Span Selection and Search for Sign-Language Video Comprehension among ASL Learners. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (Athens, Greece) *(ASSETS '22)*. Association for Computing Machinery, New York, NY, USA, Article 29, 14 pages. https://doi.org/10.1145/3517428.3544883
[49] Saad Hassan, Akhter Al Amin, Alexis Gordon, Sooyeon Lee, and Matt Huenerfauth. 2022. Design and Evaluation of Hybrid Search for American Sign Language to English Dictionaries: Making the Most of Imperfect Sign Recognition. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>New Orleans</city>, <state>LA</state>, <country>USA</country>, </conf-loc>) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 195, 13 pages. https://doi.org/10.1145/3491102.3501986
[50] Google Help. 2022. Share Clips - Youtube Help. https://support.google.com/youtube/answer/10332730?hl=en
[51] Jon Chao Hong, Mei Lien Chen, and Jian Hong Ye. 2020. Acceptance of YouTube applied to dance learning. *International Journal of Information and Education Technology* 10, 1 (2020), 7–13.
[52] Lu-Ho Hsia, Iwen Huang, and Gwo-Jen Hwang. 2016. Effects of different online peer-feedback approaches on students' performance skills, motivation and self-efficacy in a dance course. *Computers & Education* 96 (2016), 55–71.
[53] Chi-Min Hsieh and Annie Luciani. 2005. Generating dance verbs and assisting computer choreography. In *Proceedings of the 13th Annual ACM International Conference on Multimedia* (, Hilton, Singapore,) *(MULTIMEDIA '05)*. Association for Computing Machinery, New York, NY, USA, 774–782. https://doi.org/10.1145/1101149.1101314
[54] Javid Iqbal and Manjit Singh Sidhu. 2021. Augmented Reality-Based Dance Training System: A Study of Its Acceptance. In *Design, Operation and Evaluation of Mobile Communications: Second International Conference, MOBILE 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings*. Springer-Verlag, Berlin, Heidelberg, 219–228. https://doi.org/10.1007/978-3-030-77025-9_19
[55] Alexander Ivanov, Tim Au Yeung, Kathryn Blair, Kurtis Danyluk, Georgina Freeman, Marcus Friedel, Carmen Hull, Michael Yuk-Shing Hung, Sydney Pratte, and Wesley Willett. 2022. One Week in the Future: Previs Design Futuring for HCI Research. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 542, 15 pages. https://doi.org/10.1145/3491102.3517584
[56] J Jayanthi and P. Uma Maheswari. 2023. Reviving Cultural Heritage: Advancements in Pose Recognition and 3D Model Reconstruction for Indian Classical Dance Using AI and Augmented Reality Technologies. https://doi.org/10.21203/rs.3.rs-3522342/v1
[57] K. Kahol, Priyamvada Tripathi, Sethuraman Panchanathan, and T. Rikakis. 2003. Gesture segmentation in complex motion sequences. In *Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429)*, Vol. 2. IEEE, Piscataway, New Jersey, II–105. https://doi.org/10.1109/ICIP.2003.1246627
[58] Rajkumar Kannan, Frederic Andres, Fernando Ferri, and Patrizia Grifoni. 2011. Towards Multimodal Capture, Annotation and Semantic Retrieval from Performing Arts. In *Advances in Computing and Communications*, Ajith Abraham, Jaime Lloret Mauri, John F. Buford, Junichi Suzuki, and Sabu M. Thampi (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 79–88.
[59] Rajkumar Kannan, Frederic Andres, and Christian Guetl. 2010. DanVideo: an MPEG-7 authoring and retrieval system for dance videos. *Multimedia Tools and Applications* 46 (2010), 545–572.
[60] Falisha J Karpati, Chiara Giacosa, Nicholas EV Foster, Virginia B Penhune, and Krista L Hyde. 2015. Dance and the brain: a review. *Annals of the New York Academy of Sciences* 1337, 1 (2015), 140–146.
[61] Iris Kico and Fotis Liarokapis. 2020. Investigating the learning process of folk dances using mobile augmented reality. *Applied Sciences* 10, 2 (2020), 599.
[62] Salva Kirakosian, Grigoris Daskalogrigorakis, Emmanuel Maravelakis, and Katerina Mania. 2021. Near-contact Person-to-3D Character Dance Training: Comparing AR and VR for Interactive Entertainment. In *2021 IEEE Conference on Games (CoG)* (Copenhagen, Denmark). IEEE Press, Piscataway, New Jersey, 1–5. https://doi.org/10.1109/CoG52621.2021.9619037
[63] P. V. V. Kishore, K. V. V. Kumar, E. Kiran Kumar, A. S. C. S. Sastry, M. Teja Kiran, D. Anil Kumar, M. V. D. Prasad, and Lin Wu. 2018. Indian Classical Dance Action Identification and Classification with Convolutional Neural Networks. *Adv. MultiMedia* 2018 (jan 2018), 10 pages. https://doi.org/10.1155/2018/5141402
[64] H. Kuehne, H. Jhuang, E. Garrote, T. Poggio, and T. Serre. 2011. s. In *2011 International Conference on Computer Vision*. IEEE, Piscataway, New Jersey, 2556–2563. https://doi.org/10.1109/ICCV.2011.6126543
[65] Matthew Kyan, Guoyu Sun, Haiyan Li, Ling Zhong, Paisarn Muneesawang, Nan Dong, Bruce Elder, and Ling Guan. 2015. An approach to ballet dance training through ms kinect and visualization in a cave virtual reality environment. *ACM Transactions on Intelligent Systems and Technology (TIST)* 6, 2 (2015), 1–37.
[66] Rudolf Laban and Lisa Ullmann. 1971. *The mastery of movement.* ERIC, Plays, Inc., 8 Arlington Street, Boston, Mass. 02116.
[67] Rudolf von Laban. 1966. *Choreutics.* Macdonald & Evans, LTD, London.
[68] Rudolf von Laban and F. C. Lawrence. 1974. *Effort; economy in Body Movement.* Macdonald &; Evans, London.
[69] John Lansdown. 1978. The computer in choreography. *Computer* 11, 08 (1978), 19–30.
[70] François-Joseph Lapointe. 2005. Choreogenetics: The Generation of Choreographic Variants through Genetic Mutations and Selection. In *Proceedings of the 7th Annual Workshop on Genetic and Evolutionary Computation* (Washington, D.C.) *(GECCO '05)*. Association for Computing Machinery, New York, NY, USA, 366–369. https://doi.org/10.1145/1102256.1102338
[71] François-Joseph Lapointe and Martine Époque. 2005. The Dancing Genome Project: Generation of a Human-Computer Choreography Using a Genetic Algorithm. In *Proceedings of the 13th Annual ACM International Conference on Multimedia* (Hilton, Singapore) *(MULTIMEDIA '05)*. Association for Computing

Machinery, New York, NY, USA, 555–558. https://doi.org/10.1145/1101149.1101276

[72] Soohwan Lee and Kyungho Lee. 2023. CheerUp: A Real-time Ambient Visualization of Cheerleading Pose Similarity. In *Companion Proceedings of the 28th International Conference on Intelligent User Interfaces* (Sydney, NSW, Australia) *(IUI '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 72–74. https://doi.org/10.1145/3581754.3584135

[73] Aijun Li, Fang Zheng, William Byrne, Pascale Fung, Terri Kamm, Yi Liu, Zhanjiang Song, Umar Ruhi, Veera Venkataramani, and XiaoXia Chen. 2000. CASS: a phonetically transcribed corpus of mandarin spontaneous speech. In *Proc. 6th International Conference on Spoken Language Processing (ICSLP 2000)*. ISCA Speech, Beijing, China, vol. 1, 485–488. https://doi.org/10.21437/ICSLP.2000-120

[74] Zihao Li. 2021. Creativity and opportunity: how COVID-19 fosters digital dance education. *Digital Creativity* 32, 3 (2021), 188–207.

[75] Daniel A London, Ryley K Zastrow, Matthew J Gluck, and Paul J Cagle. 2021. Experience level influences users' interactions with and expectations for online surgical videos: a mixed-methods study. *The Journal of Hand Surgery* 46, 7 (2021), 560–574.

[76] Georgios Loupas, Theodora Pistola, Sotiris Diplaris, Konstantinos Ioannidis, Stefanos Vrochidis, and Ioannis Kompatsiaris. 2023. Comparison of Deep Learning Techniques for Video-Based Automatic Recognition of Greek Folk Dances. In *MultiMedia Modeling*. Springer Nature Switzerland, Gewerbestrasse 11, 6330 Cham, Switzerland, 325–336. https://doi.org/10.1007/978-3-031-27818-1_27

[77] Vittorio Marone and Ruben C Rodriguez. 2019. "What's so awesome with YouTube": Learning music with social media celebrities. *Online Journal of Communication and Media Technologies* 9, 4 (2019), e201928.

[78] Claire Monroy and Laura Wagner. 2023. Finding Structure in Modern Dance. *Cognitive Science* 47, 11 (2023), e13375.

[79] Amit Moryossef, Zifan Jiang, Mathias Müller, Sarah Ebling, and Yoav Goldberg. 2023. Linguistically Motivated Sign Language Segmentation. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, Houda Bouamor, Juan Pino, and Kalika Bali (Eds.). Association for Computational Linguistics, Singapore, 12703–12724. https://doi.org/10.18653/v1/2023.findings-emnlp.846

[80] Ruth Muelas Pérez, Rafael Sabido Solana, David Barbado Murillo, and Francisco Javier Moreno Hernández. 2014. Visual availability, balance performance and movement complexity in dancers. *Gait & Posture* 40, 4 (2014), 556–560. https://doi.org/10.1016/j.gaitpost.2014.06.021

[81] Mario Nakazawa and Andrea Paezold-Ruehl. 2009. DANCING, Dance ANd Choreography: An Intelligent Nondeterministic Generator. In *The Fifth Richard Tapia Celebration of Diversity in Computing Conference: Intellect, Initiatives, Insight, and Innovations* (Portland, Oregon) *(TAPIA '09)*. Association for Computing Machinery, New York, NY, USA, 30–34. https://doi.org/10.1145/1565799.1565807

[82] Adebunmi E. Odefunso, Esteban Garcia Bravo, and Yingjie V. Chen. 2022. Traditional African Dances Preservation Using Deep Learning Techniques. *Proc. ACM Comput. Graph. Interact. Tech.* 5, 4, Article 35 (sep 2022), 11 pages. https://doi.org/10.1145/3533608

[83] Maria Palazzi, Norah Zuniga Shaw, William Forsythe, Matthew Lewis, Beth Albright, Michael Andereck, Sucheta Bhatawadekar, Hyowon Ban, Andrew Calhoun, Jane Drozd, Joshua Fry, Melissa Quintanilha, Anna Reed, Benjamin Schroeder, Lily Skove, Ashley Thorndike, Mary Twohig, Ola Ahlqvist, Peter Chan, Noel Cressie, Stephen Turk, Jill Johnson, Christopher Roman, Elizabeth Waterhouse, Scott deLahunta, Patrick Haggard, and Alva Noe. 2009. Synchronous Objects for One Flat Thing, Reproduced. In *ACM SIGGRAPH 2009 Art Gallery* (New Orleans, Louisiana) *(SIGGRAPH '09)*. Association for Computing Machinery, New York, NY, USA, Article 37, 1 pages. https://doi.org/10.1145/1667265.1667306

[84] Costas Panagiotakis, Antonis Argyros, and Damien Michel. 2014. Temporal Segmentation and Seamless Stitching of Motion Patterns for Synthesizing Novel Animations of Periodic Dances. In *Proceedings of the 2014 22nd International Conference on Pattern Recognition (ICPR '14)*. IEEE Computer Society, USA, 1892–1897. https://doi.org/10.1109/ICPR.2014.331

[85] Yawen Pang and Yi Niu. 2023. Dance Video Motion Recognition Based on Computer Vision and Image Processing. *Applied Artificial Intelligence* 37, 1 (2023), 2226962.

[86] Mila Parrish. 2016. Toward transformation: Digital tools for online dance pedagogy. *Arts education policy review* 117, 3 (2016), 168–182.

[87] Stefan Petscharnig and Klaus Schöffmann. 2017. Deep Learning for Shot Classification in Gynecologic Surgery Videos. In *MultiMedia Modeling*, Laurent Amsaleg, Gylfi Þór Guðmundsson, Cathal Gurrin, Björn Þór Jónsson, and Shin'ichi Satoh (Eds.). Springer International Publishing, Cham, 702–713.

[88] John Sören Pettersson and Malin Wik. 2015. The Longevity of General Purpose Wizard-of-Oz Tools. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction* (Parkville, VIC, Australia) *(OzCHI '15)*. Association for Computing Machinery, New York, NY, USA, 422–426. https://doi.org/10.1145/2838739.2838825

[89] Stefano Piana, Paolo Coletta, Simone Ghisio, Radoslaw Niewiadomski, Maurizio Mancini, Roberto Sagoleo, Gualtiero Volpe, and Antonio Camurri. 2016. Towards a Multimodal Repository of Expressive Movement Qualities in Dance. In *Proceedings of the 3rd International Symposium on Movement and Computing* (Thessaloniki, GA, Greece) *(MOCO '16)*. Association for Computing Machinery, New York, NY, USA, Article 10, 8 pages. https://doi.org/10.1145/2948910.2948931

[90] Efrosini Protopapa. 2015. A Choreographer's Score: Fase, Rosas Danst Rosas, Elena's Aria, Bartók. *Dance Research Journal* 47, 2 (2015), 80–84.

[91] Cláudia Ribeiro, Rafael Kuffner, and Carla Fernandes. 2022. *10 Capturing and visualizing 3D dance data: Challenges and lessons learnt.* Taylor & Francis, Milton Park Abingdon, United Kingdom.

[92] Claudia Ribeiro, Rafael Kuffner, Carla Fernandes, and João Pereira. 2016. 3D Annotation in Contemporary Dance: Enhancing the Creation-Tool Video Annotator. In *Proceedings of the 3rd International Symposium on Movement and Computing* (Thessaloniki, GA, Greece) *(MOCO '16)*. Association for Computing Machinery, New York, NY, USA, Article 41, 4 pages. https://doi.org/10.1145/2948910.2948961

[93] Jean-Philippe Rivière, Sarah Fdili Alaoui, Baptiste Caramiaux, and Wendy E. Mackay. 2018. How Do Dancers Learn To Dance? A first-person perspective of dance acquisition by expert contemporary dancers. In *Proceedings of the 5th International Conference on Movement and Computing* (Genoa, Italy) *(MOCO '18)*. Association for Computing Machinery, New York, NY, USA, Article 6, 7 pages. https://doi.org/10.1145/3212721.3212723

[94] Jean-Philippe Rivière, Sarah Fdili Alaoui, Baptiste Caramiaux, and Wendy E. Mackay. 2018. How Do Dancers Learn To Dance? A First-Person Perspective of Dance Acquisition by Expert Contemporary Dancers. In *Proceedings of the 5th International Conference on Movement and Computing* (Genoa, Italy) *(MOCO '18)*. Association for Computing Machinery, New York, NY, USA, Article 6, 7 pages. https://doi.org/10.1145/3212721.3212723

[95] Charles Rose, Michael F Cohen, and Bobby Bodenheimer. 1998. Verbs and adverbs: Multidimensional motion interpolation. *IEEE Computer Graphics and Applications* 18, 5 (1998), 32–40.

[96] Sharon Māhealani Rowe. 2008. We dance for knowledge. *Dance Research Journal* 40, 1 (2008), 31–44.

[97] Vidhu V. Saxena, Tommy Feldt, and Mohit Goel. 2014. Augmented Telepresence as a Tool for Immersive Simulated Dancing in Experience and Learning. In *Proceedings of the 6th Indian Conference on Human-Computer Interaction* (New Delhi, India) *(IndiaHCI '14)*. Association for Computing Machinery, New York, NY, USA, 86–89. https://doi.org/10.1145/2676702.2676708

[98] Thecla Schiphorst. 1993. *A Case Study of Merce Cunningham's Use of the Lifeforms Computer Choreographic System in the Making of Trackers.* Simon Fraser University, 8888 University Drive Burnaby, B.C.. Canada. V5A 1S6.

[99] T. Schiphorst, T. Calvert, C. Lee, C. Welman, and S. Gaudet. 1990. Tools for Interaction with the Creative Process of Composition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, Washington, USA) *(CHI '90)*. Association for Computing Machinery, New York, NY, USA, 167–174. https://doi.org/10.1145/97243.97270

[100] Sherry B Shapiro. 2008. *Dance in a world of change: Reflections on globalization and cultural difference.* Human Kinetics, Champaign, IL 61820.

[101] Vikash Singh, Celine Latulipe, Erin Carroll, and Danielle Lottridge. 2011. The Choreographer's Notebook: A Video Annotation System for Dancers and Choreographers. In *Proceedings of the 8th ACM Conference on Creativity and Cognition* (Atlanta, Georgia, USA) *(C&C '11)*. Association for Computing Machinery, New York, NY, USA, 197–206. https://doi.org/10.1145/2069618.2069653

[102] Vikash Singh, Celine Latulipe, Erin Carroll, and Danielle Lottridge. 2011. The choreographer's notebook: a video annotation system for dancers and choreographers. In *Proceedings of the 8th ACM Conference on Creativity and Cognition* (Atlanta, Georgia, USA) *(C&C '11)*. Association for Computing Machinery, New York, NY, USA, 197–206. https://doi.org/10.1145/2069618.2069653

[103] Mary M Smyth. 1984. Kinesthetic communication in dance. *Dance Research Journal* 16, 2 (1984), 19–22.

[104] Asako Soga, Bin Umino, Takami Yasuda, and Shigeki Yokoi. 2006. Web3D dance composer: automatic composition of ballet sequences. In *ACM SIGGRAPH 2006 Research Posters* (Boston, Massachusetts) *(SIGGRAPH '06)*. Association for Computing Machinery, New York, NY, USA, 5–es. https://doi.org/10.1145/1179622.1179628

[105] Asako Soga, Bin Umino, Takami Yasuda, and Shigeki Yokoi. 2007. Automatic composition and simulation system for ballet sequences. *The Visual Computer* 23 (2007), 309–316.

[106] Statista. 2023. Dance exercise participants US 2022 | Statista. (Accessed on 09/12/2023).

[107] Dan Sun. 2021. Dance Training Movement Depth Information Recognition Based on Artificial Intelligence. In *2020 International Conference on Applications and Techniques in Cyber Intelligence*, Jemal H. Abawajy, Kim-Kwang Raymond Choo, Zheng Xu, and Mohammed Atiquzzaman (Eds.). Springer International Publishing, Cham, 355–360.

[108] Susan Tendy and Ivan Evans. 2008. Ice Skating. , 313 pages.

[109] American Ballet Theatre. 2018. Ballet Dictionary. https://www.abt.org/explore/learn/ballet-dictionary/

[110] Jing Tian, Xiaoqiang Yang, and Hangjun Che. 2022. Dance Movement Recognition Based on Modified GMM-Based Motion Target Detection Algorithm. *Sec. and Commun. Netw.* 2022 (jan 2022), 12 pages. https://doi.org/10.1155/2022/

6023784
[111] Himberg Tommi and Thompson Marc R. 2011. Learning and synchronising dance movements in South African songs–cross-cultural motion-capture study. *Dance Research* 29, supplement (2011), 305–328.
[112] Charlotte Veal. 2016. A choreographic notebook: Methodological developments in qualitative geographical research. *cultural geographies* 23, 2 (2016), 221–245.
[113] Elizabeth Walton and Wendy E. Mackay. 2022. Dance Transitions: What Forms of Technology Best Support Professional Dancers as They Learn New Movement Styles?. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI ’22)*. Association for Computing Machinery, New York, NY, USA, Article 223, 14 pages. https://doi.org/10.1145/3491102.3517448
[114] Meagan A Woodard. 2020. Dance: The Universal Language of Storytellers.
[115] Ming Yan and Zhe He. 2022. Dance Action Recognition Model Using Deep Learning Network in Streaming Media Environment. *Journal of Environmental and Public Health* 2022 (2022), 11 pages.
[116] Zhenyu Yang, Bin Yu, Wanmin Wu, Ross Diankov, and Ruzena Bajscy. 2006. Collaborative Dancing in Tele-Immersive Environment. In *Proceedings of the 14th ACM International Conference on Multimedia* (Santa Barbara, CA, USA) *(MM ’06)*. Association for Computing Machinery, New York, NY, USA, 723–726. https://doi.org/10.1145/1180639.1180793
[117] Andrew Yee, William M Padovano, Ida K Fox, Elspeth JR Hill, Amanda G Rowe, L Michael Brunt, Amy M Moore, Alison K Snyder-Warwick, Lorna C Kahn, Matthew D Wood, et al. 2020. Video-based learning in surgery: establishing surgeon engagement and utilization of variable-duration videos. *Annals of surgery* 272, 6 (2020), 1012–1019.
[118] Giugno Ylenia, Napolitano Salvatore, Izzo Riccardo, and Raiola Gaetano. 2013. Assessment of aerobic gymnastics by video analysis. *ASSESSMENT* 13, 2 (2013), 205–210.
[119] Tina Yu and Paul Johnson. 2003. Tour Jeté, Pirouette: Dance Choreographing by Computers. In *Genetic and Evolutionary Computation — GECCO 2003*, Erick Cantú-Paz, James A. Foster, Kalyanmoy Deb, Lawrence David Davis, Rajkumar Roy, Una-May O’Reilly, Hans-Georg Beyer, Russell Standish, Graham Kendall, Stewart Wilson, Mark Harman, Joachim Wegener, Dipankar Dasgupta, Mitch A. Potter, Alan C. Schultz, Kathryn A. Dowsland, Natasha Jonoska, and Julian Miller (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 156–157.
[120] Xianfeng Zhai. 2021. Dance movement recognition based on feature expression and attribute mining. *Complexity* 2021 (2021), 1–12.
[121] Qiushi Zhou, Cheng Cheng Chua, Jarrod Knibbe, Jorge Goncalves, and Eduardo Velloso. 2021. Dance and Choreography in HCI: A Two-Decade Retrospective. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI ’21)*. Association for Computing Machinery, New York, NY, USA, Article 262, 14 pages. https://doi.org/10.1145/3411764.3445804
[122] Äli Leijen, Ineke Lam, Liesbeth Wildschut, P. Robert-Jan Simons, and Wilfried Admiraal. 2009. Streaming video to enhance students’ reflection in dance education. *Computers & Education* 52, 1 (2009), 169–176. https://doi.org/10.1016/j.compedu.2008.07.010

A VIDEO STIMULI USED IN STUDY 2

We summarize the dance sequences shown in the six videos to the participants. The difficulty level is given in the bracket along with a short description. The descriptions are adapted from [39].

(1) Siete Loco (Advanced): In this dance move, the leader and follower begin in closed position. They perform a Siete con Coca Cola move, followed by a back-to-back Vacilala move with the leader grabbing the follower’s hand. They then perform two Exhibela moves, with the leader placing their hand on the follower’s neck on the second one.
(2) Kentucky Complicado (Advanced): In the Kentucky Complicado dance, the couple starts in a closed position and performs two Kentucky moves before initiating the Exhibela move. The leader then does an Enchufala Engancha followed by a Dile Que No con Coca Cola move, resulting in a “back to back" position. To “get out” of the position, the leader does a Side Rocks step and a Vuelta move, ending in a closed position. Finally, they perform two more Kentucky moves.
(3) Festival Prima (Improver with four performers): The dancers start in closed position and perform a sequence of three Prima moves: Prima Con La Hermana, Prima Esquipi, and Prima con Paulito. Each move is finished with a Dame step.
(4) El Duque Loco (Master): In El Duque Loco, the dancers start in closed position and perform El Duque followed by two Exhibela moves, a left turn with Engancha, and a switch to right Engancha before ending with an Enchufala.
(5) Beso Complicado (Advanced): In the beginning, the leader and follower are in a closed position and start with the first few steps of Beso. Then, on the next beat, the leader lifts their right arm, allowing the follower to move around them. Afterward, they do a Vacilala move, where the follower exits first and the leader follows. Then, they switch positions to Guapea and can choose to perform an Alarde. Finally, they end with an Enchufala, finishing in the “crossed hands ending.”
(6) Bayamo (Advanced): The dance starts with a right-to-right handhold. The leader and follower do a series of moves that include Vacilala, Dile Que No, Vacilala with an Alarde, and Enchufala with crossed hands. The leader grabs the follower’s left hand on count 6. There’s a variation of the dance called “Ponle El Cagua Uno,” which starts with a Sombrero instead of a Vacilala con la Mano, but it’s not commonly used.