Classifying Physical Strategies in Tangible Tasks: A Video-Coding Framework for Epistemic Actions

Augusto Esteves
Exact Sciences and Eng. Center
University of Madeira, Portugal
augustoeae@gmail.com

Saskia Bakker
Department of Industrial Design
Eindhoven University of Technology, Netherlands
s.bakker@tue.nl

Alissa Antle
Interactive Arts and Technology
Simon Fraser University, Canada
aantle@sfu.ca

Aaron May
Interactive Arts and Technology
Simon Fraser University, Canada
amay@sfu.ca

Jillian Warren
Interactive Arts and Technology
Simon Fraser University, Canada
jlw29@sfu.ca

Ian Oakley
Design and Human Engineering
Ulsan National Institute of Science and Technology, Republic of Korea
ian.r.oakley@gmail.com

CHI 2014, Apr 26 - May 01 2014, Toronto, ON, Canada
ACM 978-1-4503-2474-8/14/04.
http://dx.doi.org/10.1145/2559206.2581185.

Abstract
Tangible interaction is a compelling interface paradigm that elegantly merges the fluency of physical manipulation with the flexibility of digital content. However, it is currently challenging to understand the real benefits and advantages of tangible systems. To address this problem, this paper argues that we need new evaluation techniques capable of meaningfully assessing how users perform with tangible, physical objects. Working towards this aim, it presents a video-coding framework that supports the granular identification of epistemic actions (physical actions that are made to simplify cognitive work) during tangible tasks. The framework includes 20 epistemic actions, identified through a systematic literature review of 77 sources. We argue that data generated by applying this process will help us better understand epistemic behavior and, ultimately, lead to the generation of novel, grounded design insights to support physically-grounded cognitive strategies in tangible tasks.

Keywords
Epistemic Action; Tangible Interaction; Video Coding

ACM Classification Keywords
H.5.2.

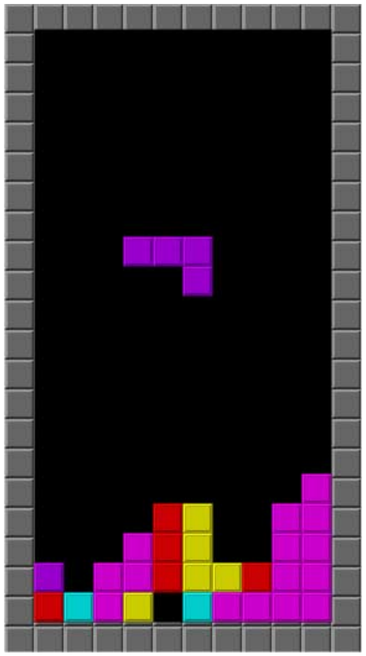

Figure 1. Kirsh et al. [9] introduced epistemic actions through the popular game of Tetris. In this game, a key example of an epistemic action includes physically rotating a descending piece so as to more easily compare its contours to those of the pieces below. This allows players to identify a suitable destination for the piece more quickly than if the rotation had been performed purely mentally.

Introduction and Related Work

In tangible interaction [11] users manipulate physical artifacts that both represent and control digital information, providing a compelling directness and physicality that is attracting increasing interest in HCI research labs. One fundamental reason for this growing popularity is that a tight coupling between the physical and the digital allows users to apply knowledge and skills gained during everyday, real-world activity to interaction with computational systems [4]. Much of the value of this idea can be grounded in the cognitive science literature on embodied and situated cognition, an emerging school of thought where the body and its interactions with the world are construed as taking a central role in human thought and experience [12, 14]. While the literature on this topic is broad, one key aspect that is particularly relevant for HCI work is the notion that users leverage external (non-mental) structures to simplify cognitive work [2]. One prominent example of this approach is Kirsh et al.'s [9] long standing distinction between epistemic actions, which are physical actions performed to simplify the mental demands of a task, and pragmatic actions, which users execute to directly bring them closer to their goals (see example in Figure 1).

However, while these theoretical ideas are powerful, meaningfully applying abstract concepts from cognitive science to practical interaction design activity remains a challenging and intricate task. We highlight two distinct and embryonic approaches in the tangible interaction literature. The first is quantitative and focuses on establishing appropriate metrics to empirically evaluate systems [1, 3]. For example, several authors have studied how users manipulate objects in a tangible system with the goal of distinguishing between pragmatic and epistemic action and better understanding the mechanisms by which physical artifacts are used to aid cognitive work. A current weakness of this work is in terms of the granularity and scope of the epistemic actions considered – these are typically either discussed in highly specific, contextualized scenarios or the diversity of possible actions is reduced to a single categorical label [e.g. 1]. This limits generalizability. At the other end of the scale, in the second approach, a number of authors have introduced design-orientated frameworks that aim to guide the creation of novel interactive systems [e.g. 5, 6, 10]. While these are valuable, they are also subjective and the insights they provide are currently based on little formal evidence. While they represent an important design resource, their validity is limited.

This paper aims to build a bridge between these two approaches by introducing a video-coding framework that enables researchers to categorize hand actions during tangible tasks according to a detailed classification scheme for epistemic actions. This scheme was developed through a systematic literature review of 77 papers, the extraction of key examples of epistemic action from these papers, and a synthesis exercise over this content. Although loosely related prior classifications exist [e.g. 8, 9], the framework presented in this paper is the first to be based a systematic review, the first to aim for a focused, fine-grained description, and the first to be specifically directed towards the development of an actionable empirical tool for capturing and expressing observed epistemic behaviors. This paper argues that the framework contributes to our understanding of how epistemic actions are used in tangible tasks and provides researchers with a tool to more systematically

Figure 2. The 20 epistemic actions part of the video-coding framework being introduced were gathered from an initial set of 335 quotes found across 77 papers, journals, and books. These pieces of work originate from a range of fields such as mathematics, cognitive science, HCI, and design, from the last three decades.

assess this complex type of behavior. The remainder of the paper describes how the framework was developed and the 20 epistemic actions that constitute it. This work-in-progress paper closes with a discussion of the future work required to assess the framework's reliability, validity, and predictive power.

Framework Development

An extensive literature review was conducted in order to capture the different possible types of epistemic action that form the video-coding framework being introduced. A set of keywords was used to execute a literature search on both Google Scholar and Science Direct. These included 'epistemic action(s)', but also 'complementary action(s)' and 'complementary strategies'. The first 60 results from each of these searches were kept for further inspection. Additionally, papers referencing seminal work (specifically [7, 9]) and including the defined keywords were also retained. Ultimately, 77 articles were obtained through this process. Each of them was then inspected for any mention of actions that could be interpreted, or were directly treated as epistemic, and quotes such as "(...) *preparing the workplace, for example, by partially sorting nuts and bolts before beginning an assembly task in order to reduce later search* (...)" (page 515, [13]) were extracted. In total 335 quotes were compiled (see Figure 2). Two of the authors worked collaboratively to create an affinity diagram that identified different clusters of epistemic actions, and of users' goals when performing such actions.

Framework Structure: Categories & Actions

Of the 335 quotes gathered, 225 were eventually retained. Discarded quotes included actions judged to have unclear epistemic value. 20 general *categories* of epistemic action were generated, each containing between one to six examples of specific epistemic actions (for easier identification when video-coding). Following is a summary of all categories of epistemic actions – a full description of the framework can be found online at:
http://www.mysecondplace.org/Framework.pdf

1. Clustering or grouping artifacts together
Examples: (i) users cluster artifacts by ordering them; (ii) users cluster artifacts into spatial groups; (iii) users cluster artifacts into piles (see Figure 3); and, (iv) users select a cluster of artifacts to carry with them.

2. Spatial arrangement of artifacts in relation to one another, the task environment, or the user(s)
Examples: (i) users place an object closer or on themselves to serve as a cue; (ii) users line up artifacts in a sequence, in their hands or in the task environment; (iii) users spatially order artifacts to match some property of the task, or the artifact (e.g. card suit); and, (iv) users spatially organize more than one artifact, in their hands and arms, or in the task environment.

3. Rearrange a representation
Examples: (i) users re-arrange an external representation or model; (ii) users spatially re-order artifacts; and, (iii) users re-order artifacts to obey some property.

4. Manipulation of an artifact
Examples: (i) users rotate, move, shake, or hold artifacts; (ii) users pass artifacts to other users; and, (iii) users divide artifacts into parts.

Figure 3. Game players often pile tokens or pieces together for quicker identification and retrieval.

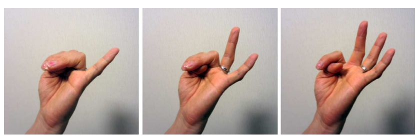

Figure 4. A common way for children to externalize a taxing internal process, such as counting, is to use their fingers as numbers.

5. Tag or annotate an artifact
Examples: (i) users color part or the entire artifact; (ii) users annotate on artifacts; (iii) users place an object or tool closer to, or on an artifact; and, (iv) users place an artifact close to or, on an object or location.

6. Using the body to externalize an internal process
Examples: (i) users use self-corrective speech; (ii) users verbalize the problem amongst them; (iii) users move their bodies or fingers (see Figure 4); and, (iv) users gesture for themselves, or to other users.

7. Annotate
Examples: (i) users write down intermediate steps; (ii) users draw a line to guide subsequent actions with an artifact or tool; and, (iii) users broadly write, type, or draw (see Figure 5).

8. Bodily marking an artifact
Examples: (i) users cue an artifact(s) in their hands or bodies; (ii) users point to an artifact(s) with their hands or tools; (iii) users mark an artifact with their finger; and, (iv) users mark an artifact(s) by bringing it closer to them.

9. Build a model or external representation
Examples: (i) users create an initial model to serve as basis for further iterations; and, (ii) users create a model or external representation.

10. Bi-manual use of two artifacts, two representations, or an artifact and a representation
Examples: (i) users rotate a representation to match another; (ii) users juxtapose two representations or two artifacts (or an artifact and a representation); and, (iii) users pick and bring two artifacts close to their eyes.

11. Divide workspace into several stations in which only a subset of actions are afforded
Examples: (i) users allocate particular tools and artifacts to areas surrounding them or other users; (ii) users move to a location with particular tools and artifacts; and, (iii) users place tools and artifacts in separate areas of the environment, not within reach of each other.

12. Rotation of the user, the problem space, or representation
Examples: (i) users get a new perspective on the problem by moving themselves around the task environment, or rotating the representation (see Figure 6).

13. Place artifact in a contrasting environment
Examples: (i) users place an artifact in contrast to a series of others; and, (ii) users differentiate artifact(s) by having them in contrasting surroundings.

14. Compare an artifact with a possible destination or other artifacts
Examples: (i) users temporarily place an artifact in one its the possible destinations; (ii) users move or rotate an artifact while comparing it with possible candidate destinations; and, (iii) users move an artifact while comparing it with other artifacts.

15. Talking or gesturing to guide and direct attention
Example: (i) users talk and gesture to better explain themselves to other users.

Figure 5. When dealing with a difficult task, users can write intermediate steps to their solution in order to alleviate possible high memory requirements, or to make their cognitive processes more reliable.

Figure 6. A tourist studies a map. It is typically easier and more reliable to physically rotate a map to match the surrounding area than to perform the equivalent rotation mentally.

16. Use of an artifact or tool to physically constraint the user, or the use of other artifacts and tools
Examples: (i) users place an artifact or tool in a way that constrains the actions performed with other artifacts; and, (ii) users rely on the task's environment to constrain the actions afforded on an artifact.

17. Clear and clean clutter
Examples: (i) users move a group of artifacts or tools out of focus or out of sight; (ii) users move an artifact(s) or tool(s) to their original placements in the beginning of the task; and, (iii) users remove an artifact(s) or tool(s) from the task environment.

18. Artifact trial-and-error positioning
Example: (i) users try to find an artifact's placement through plain trial-and-error.

19. Shuffle artifacts
Example: (i) users randomly shuffle artifacts.

20. Test the state or response of a system, model, environment, or other user
Examples: (i) users move an artifact to test other users or the system's response; (ii) users deploy an artifact to test the system's state or response; (iii) users use their bodies to test the system's state; (iv) users run a built model to better understand the system being modeled; and, (v) users test a system by operating it.

Framework Use: Instructions and Method

The framework introduced in this paper is an extension to Antle et al's hand events video-coding framework [1]. In Antle's framework, an action is classified as a *direct placement*, an *indirect placement*, or as *exploratory*. It was introduced in the context of a jigsaw puzzle task and, in this scenario, a direct placement represents situations where users already know where to place a puzzle piece before picking it up, leading to a direct transition between acquiring a piece and correctly placing it. Indirect placements represent similar outcomes but describe situations in which users are not initially certain of where to position the pieces they pick up – they translate or rotate the piece while searching for its correct destination. And finally, exploratory events represent those actions that do not finish with a piece in its correct final position.

Researchers applying the framework introduced in this paper should first categorize actions as being either direct placements, indirect placements, or exploratory. After this process is complete, all actions that were coded as either indirect placement or exploratory should be assessed for epistemic activity (i.e. do they match any of the 20 epistemic categories in the framework). It is worth noting that each indirect placement or exploratory action may contain more than one epistemic action. For example, a user may pick up a puzzle piece, compare it with possible positions in the jigsaw puzzle (epistemic action #14) and, after failing to find a suitable destination place it in a cluster of similarly colored pieces (epistemic action #1). For a more comprehensive graphical workflow of how to video-code using the framework, please consult: *www.mysecondplace.org/Coding-flowchart.pdf*

Future Work & Conclusion

The immediate follow-up to the work presented in this paper is a series of studies that demonstrate the usefulness and value of the framework as an analytic tool. These will include studies of users performing a range of tangible tasks, work that may entail adaptation of the framework. An early goal will be to consolidate and validate the categories introduced in this paper. We will also seek to assess the framework's internal validity (and predictive power) with studies that support comparing or contrasting its results with predicted performance on established metrics (e.g. time, errors, workload, or measures of relevant cognitive ability such as spatial pre-tests). Furthermore, its reliability and external validity will be determined by evaluating its consistency over a diverse range of raters, tasks and participants. In summary, this paper provides ongoing work towards capturing granular data about epistemic activity. This is a concrete and valuable step towards understanding the real benefits of interaction with tangible systems and, ultimately, generating practical design knowledge that can guide creation and improve user experience in future tangible systems.

References

[1] Antle, A. N. and Wang, S. Comparing motor-cognitive strategies for spatial problem solving with tangible and multi-touch interfaces. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, ACM (2013), 65-72.

[2] Clark, A. *Being There: Putting Brain, Body, and World Together Again*. MIT Press, 1998.

[3] Esteves, A., Hoven, E. v. d. and Oakley, I. Physical games or digital games?: comparing support for mental projection in tangible and virtual representations of a problem-solving task. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, ACM (2013), 167-174.

[4] Hornecker, E. Beyond affordance: tangibles' hybrid nature. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*, ACM (2012), 175-182.

[5] Hornecker, E. and Buur, J. Getting a grip on tangible interaction: a framework on physical space and social interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2006), 437-446.

[6] Jacob, R. J. K., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T. and Zigelbaum, J. Reality-based interaction: a framework for post-WIMP interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2008), 201-210.

[7] Kirsh, D. Complementary strategies: Why we use our hands when we think. In *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (1985), 212-217.

[8] Kirsh, D. The intelligent use of space. *Artificial Intelligence*, *73*, 1-2 (1995), 31-68.

[9] Kirsh, D. and Maglio, P. On Distinguishing Epistemic from Pragmatic Actions. *Cognitive Science*, *18*, 4 (1994), 513–549.

[10] Klemmer, S. R., Hartmann, B. and Takayama, L. How bodies matter: five themes for interaction design. In *Proceedings of the 6th conference on Designing Interactive systems*, ACM (2006), 140-149.

[11] Shaer, O. and Hornecker, E. Tangible User Interfaces: Past, Present, and Future Directions. *Foundations and Trends in Human-Computer Interaction*, *3* (2010), 1-137.

[12] Shapiro, L. *Embodied Cognition (New Problems of Philosophy)*. Routledge, 2011.

[13] Tarjan, R. E. Amortized Computational Complexity. *SIAM. J. on Algebraic and Discrete Methods*, *6*, 2 (1985), 306-318.

[14] Wilson, M. Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*, 4 (2002).