Physical Games or Digital Games? Comparing Support for Mental Projection in Tangible and Virtual Representations of a Problem-Solving Task

Augusto Esteves[1], Elise van den Hoven[2,3], Ian Oakley[1,4]
[1]Exact Sciences and Engineering Center, University of Madeira, Funchal, Portugal
[2]Design, Architecture & Building Faculty, University of Technology, Sydney, Australia
[3]Industrial Design Department, Eindhoven University of Technology, Eindhoven, the Netherlands
[4]Madeira Interactive Technologies Institute, Funchal, Portugal
augustoeae@gmail.com; elise.vandenhoven@uts.edu.au; ian.r.oakley@gmail.com

ABSTRACT
This paper explores how different interfaces to a problem-solving task affect how users perform it. Specifically, it focuses on a customized version of the game of Four-in-a-row and compares play on a physical, tangible game board with that conducted in mouse and touch-screen driven virtual versions. This is achieved through a repeated measures study involving a total of 36 participants and which explicitly assesses aspects of cognitive work through measures of time task, subjective workload, the projection of mental constructs onto external structures and the occurrence of explanatory epistemic actions. The results highlight the relevance of projection and epistemic action to this problem-solving task and suggest that the different interface forms afford instantiation of these activities in different ways. The tangible version of the system supports the most rapid execution of these actions and future work on this topic should explore the unique advantages of tangible interfaces in supporting epistemic actions.

Author Keywords
Tangible interaction; problem-solving; embodied cognition; mental projection; epistemic actions.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Human factors; Design.

INTRODUCTION
According to proponents of embodied cognition [1] the workings of the mind can be best understood, not by looking solely at the brain, but by also considering the body and its interaction with the surrounding environment. This position has direct relevance to work on tangible interaction, a user interface paradigm that couples physical objects with virtual content based on the idea that such a link is evocative, familiar [4] and capable of transferring users' real world understandings and skills to the digital domain [9]. Essentially, by engaging users through bodily actions and physical manipulations, tangible interfaces are expected to facilitate what is known as *tangible thinking* [23] – using the environment to aid cognition. While this concept is compelling (and authors have argued it is poorly supported by purely virtual systems [24]) there is still a lack of firm empirical evidence demonstrating the concrete advantages of designing user interfaces using physical, tangible technologies and approaches [2, 5, 6, 7, 17, 20].

This paper argues that determining the nature and form that such evidence needs to take is a key current issue in the field of tangible interaction; empirical findings and the understandings they enable will be crucial aspect in the process of maturing the field. Most current research efforts in tangible interaction develop valuable new knowledge by focusing on either: the development and description of novel tangible systems, making contributions in innovative hardware and software solutions [e.g. 16]; or on highly specific application areas such as the effects of tangibility on children's development and performance [19, 29, 30]. Attempting to address more general questions, a piecemeal body of research has compared the effects of relying on tangible or graphical representations on user experience. While providing valuable insights, this work suffers from being highly specific [e.g. 15, 18], 'unfair' in comparisons between interface paradigms [e.g. 24] or influenced by the novelty effect of introducing tangible systems [30].

This paper argues that, in order to effectively challenge traditional interface paradigms, an improved understanding of the benefits of tangible systems needs be established. An important aspect of achieving this is rigorous, repeatable and equivalent comparisons between interaction paradigms. Moving towards this objective, this paper presents a comparative study between three different interfaces to a problem-solving task – a game of Four-in-a-row (see Figure 1). In this game, players take turns dropping colored disks

TEI 2013, Feb 10-13, 2013, Barcelona, Spain.

in a vertical grid, with the goal of connecting four disks of the same color either horizontally, vertically, or diagonally. Two of the interfaces studied in this work show both the grid and the disks in a graphical display, differing only in the input method: mouse versus direct touch. A third interface is composed of an augmented physical game-board and real disks. Careful design of these three systems was used to ensure interactions were functionally equivalent and a user study looking at both quantitative and qualitative performance was conducted. The results highlight key similarities and differences in performance when using the three interface paradigms and suggest directions for future work on this topic to pursue.

RELATED WORK

This review covers two distinct topics. Firstly, tangible systems that support problem-solving tasks and the challenges in empirically establishing their worth. Secondly, the theories of embodied cognition [1]. These topics ground and focus the study presented in this paper.

Problem-Solving and Tangible Interaction

The external representation of a problem profoundly impacts the strategies employed to solve it – in turn, this directly and substantially affects performance [25]. Zhang [31] presents an early example of this by comparing user performance in a pair of conceptually isomorphic games: Tic-Tac-Toe and the 'Game of Fifteen', in which two players take turns selecting numbers from 1 to 9 with the goal of being the first to select three numbers that add up to 15. Zhang's work comprehensively demonstrated that users perform much slower when playing the Game of Fifteen despite the fact it is, to all intents and purposes, logically the same game as Tic-Tac-Toe.

Subsequent work in the field of tangible interaction has explored a range of more applied problem-solving tasks. For example, Urp [26] supporting urban planning tasks by allowing users to analyze buildings' shadows, proximity, or wind paths by manipulating physical objects. Similarly, Illuminating Clay [22] let users explore a landscape by deforming clay with their hands while Senseboard [10] enabled users to organize and manipulate abstract pieces of information by grouping physical pucks on a vertical grid. These examples demonstrate the general idea that tangible systems afford offloading cognitive work associated with problem-solving activities directly onto an interface [1], and that availability of a meaningful physical representation of a problem space can improve user performance.

While this notion is theoretically sound, there is little work providing empirical evidence that tangible interaction paradigms are more effective that traditional interfaces in problem-solving tasks. Indeed, most evaluations of tangible interfaces are formative in character – they seek to isolate appropriate characteristics to inform system design. While valuable, such studies are rarely critical of the design rationale or fundamental value of tangibility [30]. On the other hand, work that has attempted to conduct strict empirical investigations into the value of tangible systems has attracted a range of methodological criticisms [29, 30]. These include issues with the limited scope and generalizability of findings, as in Patten and Ishii's [21] comparison of the use of space to organize information in graphical and tangible interfaces, or in Marshall et al.'s [18] discussion of the effects of using tangible versus graphical simulations of physical systems on adults' discovery learning. Other common issues include the fact that the benefits credited to tangible interaction can be partly attributed to either novelty effects [30] or due to advantages inherent in enabling – but non-tangible – technologies such as multi-touch displays (e.g. bi-manual input) [11]. Finally, in comparisons between tangible and non-tangible systems, it can be challenging to ensure equivalence of the interfaces. For instance, in Soute et al.'s [24], investigation of tangible and virtual game objects the physical implementation introduced a range of novel functionality that was simply absent in the purely digital version.

Mental Projection and Epistemic Actions

Theoretical constructs under the banner of embodied and situated cognition fit well with the ideas underlying tangible interaction - they advocate explanations of thought that include the body and the physical and cultural constraints present in the surrounding environment [1]. One recently proposed theory introduces the concept of *projection* [14], focusing on the use external resources in the environment to simplify thinking and problem-solving tasks.

According to this work, projection is a key part of a cyclical process of problem-solving in which users act, observe the result of their actions and consider their next action [13]. In this cycle, projection sits between perception, which refers to what is sensed of the real world, and imagination, which refers to entirely mental constructs. Between these poles, projection refers to mental augmentations of reality that are anchored and grounded on perceived external structures [14]. For example, projection is the process that occurs when a person looks at a piece on a chessboard and is able to visualize the possible (or even the good vs bad) moves.

In this framing, when people finally act, they externalize a structure that is initially mental. This action can serve two distinct purposes. A *pragmatic action* will directly address the problem at hand (e.g. moving a piece to a new position) [15] whereas an *epistemic action* can serve several purposes. One is lowering the cognitive cost of projecting by instantiating some of the content in the real world (e.g. lifting a piece from the chessboard to better understand the impact of moving it). A second is to nurture additional projections (e.g. hovering a piece over a possible future location on the board to better envision additional moves) [14]. Following this logic, this paper argues that a representation of a problem that allows for a faster cycle of projection-action-projection will be less cognitively taxing and support increased user performance. It conducts a study of comparing one tangible and two non-tangible versions of

Figure 1. A standard game of Four-in-a-row [28].

the same game-based problem-solving task in order to explore the veracity of this claim.

METHOD

The goal of this paper is to provide concrete results relating to the claimed cognitive benefits of dealing with a problem-solving task through a tangible representation [23]. This is done by comparing the users' performance across three identical interfaces: a physical game set and two graphical displays in which users interact with game tokens through either a touch screen or a traditional mouse input device.

Experimental Design and Participants

The study followed a within subjects repeated measures design based on three interface conditions: tangible, touch and mouse. In total, there were 36 participants, 22 males and 14 females. 19 participants were from Europe, 14 from Asia, two from North America and one from South America. Their ages ranged from 16 to 34 (M = 24, SD = 4.16), and with the exception of one, all participants were students at local universities. Of the 36 participants, only four had never played Four-in-a-row before.

The participants completed the study in groups of three (for a total of 12 sessions) and also completed a total of three game sessions, one using each of the interface conditions. To mitigate potential practice or fatigue effects, the order in which the conditions were experienced was fully balanced – two groups completed each of the six possible order conditions. All participants received compensation in the form of a 5€ voucher valid across a range of stores and service providers. Success at the game was also rewarded – the participant who won most games in each group received an additional 5€ voucher, while the top three participants in the whole study received a further 10€ voucher.

Materials

Game Mechanics and Interactive Feedback

Four-in-a-row is a board game where two players take turns dropping colored disks in a 7x6 vertical grid (see Figure 1). The objective is to be the first player to connect four disks of the same color in either a horizontal, vertical, or diagonal line. Disks are dropped into columns from the top, meaning that the gradually changing accumulation of the disks in the different columns is an important game play element. For the purposes of this study, the Four-in-a-row game mechanics were altered. Essentially, a third player (basically a third piece color) was added, ensuring more complex changes took place between each player's moves (the introduction of two rather than one new piece). This also ensured that no users had specific prior exposure to the game dynamics, as they were as least partially novel.

Interactive feedback was also introduced to the game. This took the form of highlighting in response to exploratory gestures with the game pieces. Essentially, if participants positioned a game token at the top of one of the game board columns for a dwell period in excess of one second, they were presented with appropriately colored visual feedback indicating the position the disk would reach if dropped (see Figure 2). We termed this feature *hovering* feedback, and it was intended to provide information on the board's possible future states prior to making an actual move in the game.

System implementation

Three versions of the game were produced to support tangible, touch and mouse interaction. Each featured a 7x6 grid of holes with a total visible size of 26x24cm. Each used game disks of 3cm in diameter (0.5cm thick for the tangible version) that could be moved directly above the game board's columns to receive the hovering feedback and/or to be dropped into place. The mouse and touch versions used fully graphical interfaces developed using the Processing programming language and displayed on a small portion of a 120cm vertical flat screen. Mouse input was provided via a standard peripheral attached to the computer driving this display, while touch input was achieved via a SMART Board Interactive Display Overlay placed in front of the screen. In both these interfaces, simply clicking the mouse or touching the screen caused a drag-able icon of a game disk to appear under the cursor (or finger). This could be positioned directly above the board to gain access to the hover feedback or released there to add a piece to the game.

The tangible version was based on a physical game board into which physical disks were placed. The hover feedback was realized via two vertically stacked photo interrupters mounted on top of each of the columns (14 sensors in total). Placing a physical token in between the top emitter and sensor triggered the hovering feature (see Figure 2), while

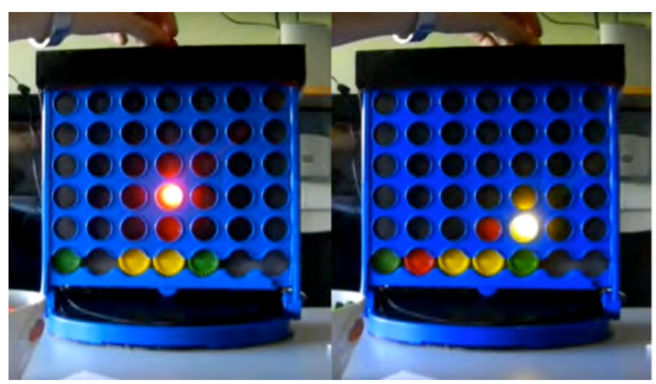

Figure 2. The tangible version of the game showing the hovering feature for a red and yellow disk (from left to right).

an interruption of the bottom sensor indicated a disk drop. Each of the bottom sensors was located 0.5cm above the game board, with the top sensors located at 1.5cm.

Graphical feedback for the hover event was enabled by placing a diffuser screen (Rosco Grey) and seven strips of digitally addressable RGB LEDs behind the board (so that there was one LED for game-board hole). All electronics were connected to an Arduino Mega microprocessor that monitored input and displayed the feedback. This construction ensured a bright, responsive display and that participants were only able to see the board from one side (as in the case of the two other versions of the system).

Procedure

In each session of the study a group of three participants played three games of Four-in-a-row against each other, one game in each of the three interfaces. Sessions commenced with a brief introduction explaining the condition sequence, game rules and compensation structure, followed by the assignment of each participant to a disk color for the duration of the study (red, yellow or green).

The experimental interfaces were all presented in the same small and otherwise empty office. Each of the three games followed an identical structure: the three participants were invited to interact informally with the interface (max. five minutes) placing disks and becoming acquainted with the hovering feature. They were then asked to move to an adjacent room, where chairs and snacks were provided. Whilst there, they were instructed not talk about the game. Participants entered the game room individually in order to make their moves, ensuring that their epistemic actions were private. The first player to move was always randomly selected but the sequence of players was always the same: red, yellow, and then green. Information reminding players of this sequence was prominently displayed in both game and waiting rooms. Both between individual turns and at the end of each game, participants completed a range of subjective measures, as detailed in the following section.

Measures

In addition to game play results, the metrics used were:

Time to play: As used in similar problems [e.g. 14], this metric is defined as the total amount of time participants take to complete their turns. The start point of this period was calculated by equipping all three versions of the game with a face recognition system consisting of a standard webcam and the Processing OpenCV computer vision library. When a participant faced the game board, this event

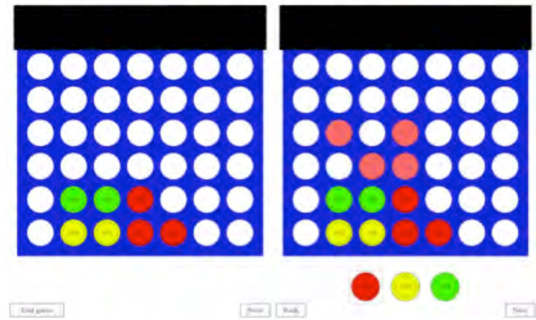

Figure 3. Screenshot of the Four-in-a-row tablet application. On the left, the board is updated with the last play (a red disk). On the right, the reasoning is given for making that play.

was recognized, a sound played and the initial time logged. The period ended when the participant dropped a disk into the grid, as detected by the game software. In order to ensure that the time taken to pick up or select a disk did not influence this measurement, participants in the tangible version started their turn with a disk already in hand. Similarly, participants in the two graphical versions "picked up" a disc simply by clicking the mouse or touching the screen, irrespective of where these events occurred.

Mental projection: In the third and final game of each session, and directly after completing each turn, participants used a tablet application to explain and justify their moves. This was achieved via a custom Android app that showed a Four-in-a-row game board and enabled them to tap grid cells to illustrate not only the current state of the game, but also the potential moves they considered whilst planning their play (see Figure 3). This application ran on 10.1" Android tablet, and was developed using the Processing programming language. The application logged two key data points: the number of candidate positions they considered for their move and the number of possible opponent responses they considered.

Epistemic actions: Epistemic actions relating to pointing were recorded in each of the interfaces both automatically, by recording when the hovering feature was triggered, and through video analysis (two observers, with a high inter-rater reliability – a Kappa of 0.7012). In this latter case, epistemic actions took the form of pointing gestures at or in front of the game board. These were divided into those made with or without the game disk (e.g. see Figure 4).

Subjective Workload: Each participant completed the NASA TLX, Hart and Staveland's six-item workload questionnaire [8], at the end of each game.

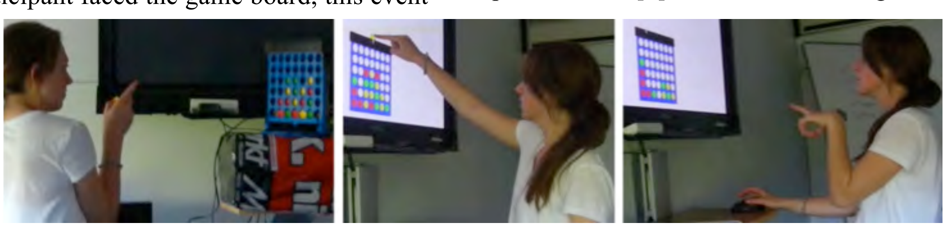

Figure 4. A participant playing Four-in-a-row in each of the interfaces (tangible, touch, and mouse). In the tangible and mouse interfaces the participant is performing a pointing action without a disk, while in the touch interface the participant triggers the hovering feature.

	Tangible	Touch	Mouse
Time to play	17.85 (5.81)	22.89 (12.3)	27.3 (13.7)

Table 1. Mean time to play in seconds according to each game interface (12 games per interface). Standard deviation in brackets.

	Tangible	Touch	Mouse
Own	6.25 (4.35)	7.50 (3.78)	13.1 (7.95)
Opponents	3.08 (3.48)	1 (0.63)	3.08 (4.12)

Table 2. Mental projection: self reported mean number of moves considered prior to play (four games per interface – the tablet application was only used in the last game of each session). Standard deviation in brackets.

	Tangible	Touch	Mouse
Hovering feature	5.37 (4.46)	2.41 (2.45)	5.56 (4.64)
With disk	1.11 (1.68)	0.56 (0.78)	1.56 (2.15)
Without disk	0.50 (1.20)	2.22 (2.05)	0.28 (0.75)

Table 3. Mean occurrence rates per game for epistemic actions recorded with the hovering feature and obtained through video analysis. Standard deviation in brackets.

RESULTS

The experimental results are now presented. Unless otherwise noted, all analyses were conducted as repeated measures one-way ANOVAs over the three experimental conditions (Tangible, Touch and Mouse). Greenhouse-Geisser corrections were used if required, and all *post-hoc* comparisons were t-tests with Bonferroni corrections.

Time to Play: The mean time to play across all interfaces is presented in Table 1. Outliers resulting from problems (e.g. jammed disks, the system failing to detect a new turn) with the tangible version of the game were removed prior to analysis. A relevant trend was found in this data (F (2, 52) = 8.202, p = 0.001) and subsequent pair-wise differences were revealed between the tangible and mouse interfaces (p = 0.004), but not between tangible and touch (p = 0.160) nor, although there was an observable trend, the mouse and touch conditions (p = 0.067).

Mental Projection: The mean results from the data recorded with the tablet application are presented in Table 2. Data was only considered after each player had made two moves, to ensure some degree of game complexity. The number of own moves considered varied significantly (independent samples ANOVA, F (2, 27) = 4.19, p = 0.026) and pair-wise comparisons showed the significant changes to be between tangible and mouse (F (11, 11) = 3.34, p = 0.029), and touch and mouse interfaces (F (11, 5) = 4.86, p = 0.047) but not between tangible and touch (p = 0.355). There were no statistically significant differences in the number of opponents' moves recorded (p = 0.430).

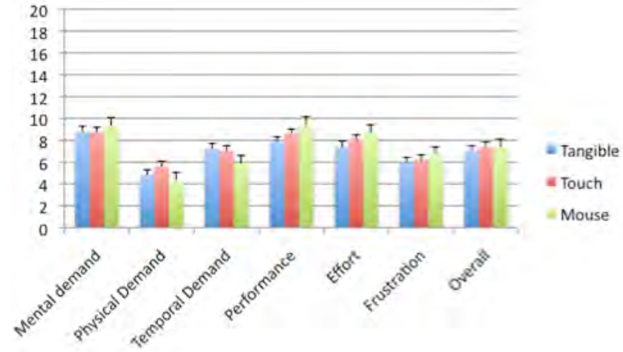

Figure 5. TLX data per condition (0-20 range). Lower scores represent lower perceived workload. Standard deviation in bars.

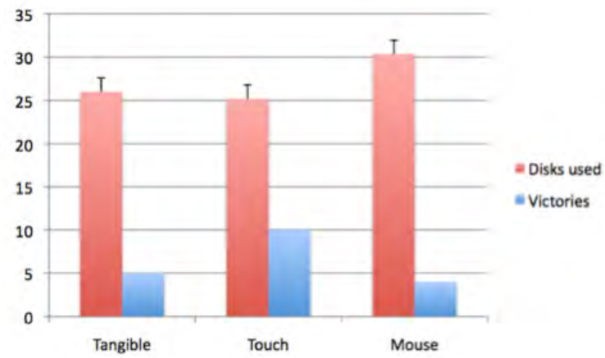

Figure 6. General game results: mean disks used in red (of a maximum of 42 per game), and number of games ending in a victory in blue (of a total of 12 games per interface). Standard deviation in bars.

Epistemic Actions: The mean results for epistemic actions performed are presented in Table 3. These are divided into those that relied on the hovering feature and those by pointing with and without a disk in hand. Data from the hovering feature was only considered when the feedback light was on for at least one second. Significant trends were observed for the use of the hovering feature (F (2, 52) = 8.772, p = 0.001) and the gestures without disks (F (1.220, 20.74) = 20.1, p < 0.001), but not for gestures with disks (F (1.173, 19.936) = 1.740, p = 0.204). For both of these two trends pair-wise comparisons revealed differences between tangible and touch, and touch and mouse conditions (both at p < 0.005 or lower). No differences between the tangible and mouse conditions were found (p > 0.489). Finally, all participants were observed performing epistemic actions. Across a full session the minimum number of such actions performed by a participant was three, the maximum 41.

Subjective Workload: The mean results from the TLX workload questionnaire are presented in Figure 5. Analyses were conducted on overall workload and each individual scale. No significant differences were observed (p > 0.177 in all cases).

Game play Statistics: The mean results for number of both disks used and victories per interface are presented in Figure 6. These show that in average, 26 (SD = 10.82) disks were used per game in the tangible, 25.2 (*SD* = 9.22) in the touch, and 30.33 (*SD* = 7.57) in the mouse version of the interface (a full board contains 42 disks). Additionally, five games ended in a victory in the tangible version (seven draws), 10 in the touch version (two draws), and four in the mouse version (eight draws). No participants won all three games, and only on one occasion did a single participant win twice in a session. Finally, 47.37% of the participants made more epistemic actions than their opponents when winning a game, and 43.75% took more time to play.

DISCUSSION AND CONCLUSIONS

This paper is grounded on Kirsh's work on mental projection [14] that investigated the time that players took to make moves in three different representations of the game of Tic-Tac-Toe. By showing that users played faster when game materials and elements were visible (such as the game board or the Xs and Os marked on that board), Kirsh demonstrated the importance of having *out-of-the-mind* structures on which to anchor cognition. Moving beyond these findings, this paper's goal was to study if a physical, tangible representation of a problem could serve as a better anchor for users' cognitive endeavors than (as otherwise as similar as possible) purely graphical counterparts.

The basic experimental results are ambiguous on the benefits of tangible interaction style over the two virtual systems studied. The time to play data showed one significant difference – the tangible interface improved over mouse input but not touch-screen. However, this can be partly explained by prior authors' assertions that dragging objects with a mouse is more time consuming than performing the same action through the more direct input methods available in tangible and touch interaction styles [3]. Furthermore, while there are noticeable differences in the data relating to mental projection of other player's moves, it is worth noting that this comparison was subject to the influence of individual differences (e.g. it was not within-subjects). Furthermore, an ANOVA on the total number of mental projections did not yield a significant result ($p = 0.068$). An alternative explanation for the increased rates of projection reported in the mouse interface is simply that the additional time moving pieces took afforded more opportunities to think about the game-board.

These results cast doubt on the value of physical representations in problem-solving tasks. However, before dismissing them, it is worth extending this discussion to include a more in-depth consideration of the cognitive work involved in problem solving. As discussed previously, epistemic actions allow users to reduce the cognitive cost of maintaining or extending mental projections, as they allow users to externalize aspects of these artifacts – to make parts of them real [14, 15]. This is a key factor underpinning the notion of thinking with things, a process that Kirsh characterizes as: "knowing what you are thinking by seeing what you are saying" [14]. A modern example of the importance of epistemic actions in problem-solving tasks comes from an anecdote about the Chess game available for the Microsoft PixelSense tabletop computer. An early version of this game only permitted valid moves but, in response to user feedback, an update was released that allowed for unconstrained and exploratory moves such as those that can be natively performed on a real chess board.

Reflecting the importance of epistemic actions, this paper choose to focus on three specific behaviors: the hovering feature, as it enabled an explicit and automatically recordable epistemic action; and pointing or touching the game board (both with and without a disk) a general epistemic technique that is reported to help users focus attention through symbolic marking [12]. Variations in the occurrence rates of these three actions in the study were complex. In the case of disk-hovers, rates were down in the touch interface while for unencumbered pointing, they were up. Conversely, no differences were observed for gestures when a user was holding a disk. These data stand in contrast to claims, typically based on the ease with which physical tokens can be grasped and manipulated, that tangible interaction is more suited to supporting epistemic actions than graphical interfaces [e.g. 21]. Rather it highlights that epistemic actions are readily achievable in the digital domain, a suggestion supported by Kirsh's seminal work introducing the concept using a traditional (and purely virtual) version of the Tetris video game [15]. Indeed, other authors have remarked on the diverse and flexible mechanisms by which people achieve and employ epistemic actions [e.g. 27].

This perspective helps explain the fact that while some of the epistemic actions were more commonly performed in the tangible interface (e.g. twice the number of hovering actions were observed compared to the touch interface), others were equivalently or less frequently performed. This suggests that people are strongly disposed to using epistemic actions and highly effective at taking advantage of whatever resources are available and optimal to realize them. Furthermore, it is clear that the different interfaces and representations of the problem afforded different actions – gesturing with a coin in hand was arguably simpler in the tangible interface than in the touch interface, where one would first have to come in contact with the touch screen before the gesture could be achieved.

Another key question of interest is whether mental projection and epistemic actions positively influenced user performance across the study as a whole. In order to fully consider this issue, it is worth highlighting the novel aspects of the game play. Introducing a third player (and consequently an additional color of disk) to the game while maintaining the board size and winning criteria (a line of four in length) made it substantially more challenging. This effect can be seen in the relatively even distribution of wins across the study – normally occurring individual differences in skill levels at the original Four-in-a-row game had little influence on outcomes in the experimental task. Indeed,

	Tangible	Touch	Mouse
Mean Epistemic Actions	6.98	5.19	7.4
Total Drawn Games	7	2	8

Table 4. Mean occurrence rates of epistemic actions per game vs number of games drawn for the three interfaces.

only one player managed to win more than one game. In light of this, a draw was interpreted as representing a balanced game, where participants successfully predicted and prevented their opponents' plans. Although too speculative to be subjected to formal statistical analysis, Table 4 shows that the number of draws varied with the number of epistemic actions. This tentative relationship suggests that epistemic actions played a valuable role in helping participants understand the state of the game.

This idea is further supported by an analysis of the activities of winning participants. Essentially in 47.37% of wins, the victor was the participant who performed the greater number of epistemic actions. Considering the chance rate of adopting this position is equal among players, or 33% in the game studied here, this fact suggests users who were performing more epistemic actions were more likely to win regardless of the interface they were using. This assertion provides tacit support for the tangible interface - although the total number of epistemic actions did not vary from interface to interface, the time data suggests they were performed substantially faster in the tangible game – 6.98 actions in just 17.85 seconds versus 7.4 in 27.32 seconds while operating the mouse driven interface. This result suggests that a key advantage of tangible interfaces may be that they have the potential to support rapid execution of useful and informative epistemic actions.

To conclude, meaningfully comparing the influence of a physical interface on a problem-solving task is challenging. Tangible interaction naturally lends itself to 'unfair' comparisons as it offers interface features that cannot be effectively matched to graphical counterparts. As such, the main concern when developing the three versions of Four-in-a-row for this study was to ensure they behaved and responded consistently and equivalently. A range of techniques were used to achieve this including removing the impact of disk selection by having participants start their turns at the tangible interface with a disk already in hand and simply clicking the mouse or touching the screen to summon a drag-able disk in the graphical conditions. Furthermore, participants were constrained to release the disks in the same set of valid positions at the top of the game board in all three interfaces. The study was also explicitly designed to promote and isolate a clear and observable projection-action-projection cycle by having participants leave the game room in-between turns. This ensured they faced a substantially evolved board state afresh each and every time they needed to play a piece.

This paper argues that only with such measures in place is it possible to meaningfully compare performance and attempt to understand the effects of the physicality of a representation on the mental effort required to solve a problem. The results of the study in this paper were informative initial steps towards this goal, but many questions remain unanswered. One of the most important relates to the fact that mental projection is not a free process [14]. In the context of the game studied here it requires the anchoring of *imagined* game disks to the *real* game board. Exploring the tradeoff between this cost and the benefits it confers would be a good topic for future work. Furthermore, it has been reported that the usefulness of anchors depends on both a person's visualization abilities and the overall complexity of the problem at hand [14]. Exploring how tangible representations influence these thresholds, essentially relating to the point at which mental projection and epistemic actions become *profitable* for particular users would likely be fruitful and interesting.

The work presented in this paper also highlights the difficulty of formally demonstrating value in the tangible interaction paradigm. It does so by equipping itself with an appropriate toolbox composed of a theoretical proposition, a carefully designed set of alternative systems and a hypothesis for how these will interact. However, although the results hint at advantages of the tangible approach, few direct effects were observed. The study did successfully highlight the importance of the theoretical ideas on which it was based – mental projection and epistemic actions – and the results indicate that the form in which users will instantiate these concepts systematically varies from interface to interface. One valuable outcome is therefore that future work on tangible systems should attempt to understand and design projection techniques that are well matched to instantiation in physical artifacts.

In closing, it is important to note that this paper does not argue that tangible systems should be limited to representations replicable in graphical counterparts. The value of physical interfaces should, ultimately, be assessed on their intrinsic merit. However, comparisons can allow us to better understand when tangibility provides benefits and also yield insights into the precise form and nature of those properties. This is the perspective adopted in this paper and, as the field of tangible interaction matures, we firmly believe that continued efforts to meaningfully compare tangible systems against other interface paradigms will help shed valuable light on the advantages tangibility has to offer and elucidate the conceptual challenges it entails.

ACKNOWLEDGMENTS

A very special thanks to: the members from the UCE group at the Industrial Design department in TU/e; Jayant Venkatanathan and Saskia Bakker; and all our study participants. This work was supported by the Portuguese Foundation of Science and Technology (FCT), under the SFRH/BD70203/2010 PhD scholarship.

REFERENCES

1. Anderson, M.L. 2003. Embodied cognition: a field guide. *Artif. Intell.* 149, 1 (September), 91-130.
2. Antle, A.N. 2007. The CTI framework: informing the design of tangible systems for children. In *Proc. of TEI '07.* ACM, NY, USA, 195-202.
3. Antle, A.N. 2012. Exploring how children use their hands to think: An embodied interactional analysis. *Behaviour and Information Technology*, in press.
4. Billinghurst, M., Grasset, R., and Looser, J. 2005. Designing augmented reality interfaces. *SIGGRAPH Comput. Graph*. 39, 1 (February), 17-22.
5. Fernaeus, Y., Tholander, J., and Jonsson, M. 2008. Towards a new set of ideals: consequences of the practice turn in tangible interaction. In *Proc. of TEI '08*. ACM, NY, USA, 223-230.
6. Fishkin, K.P., Gujar, A., Harrison, B.L., Moran, T.P. and Want, R. 2000. Embodied user interfaces for really direct manipulation. *Commun*. ACM 43, 9 (September), 74-80.
7. Gaver, W.W. 1991. Technology affordances. In *Proc. of CHI '91*, Scott P. Robertson, Gary M. Olson, and Judith S. Olson (Eds.). ACM, NY, USA, 79-84.
8. Hart, S.G. and Staveland, L.E. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Human Mental Workload*, Hancock and Meshkati (Eds.*)*.
9. Hornecker, E. 2012. Beyond affordance: tangibles' hybrid nature. In *Proc. of TEI '12*, ACM, NY, USA, 175-182.
10. Jacob, R.J.K., Ishii, H., Pangaro, G., and Patten, J. 2002. A tangible interface for organizing information using a grid. In *Proc. of CHI '02*. ACM, NY, USA, 339-346.
11. Kirk, D., Sellen, A., Taylor, S., Villar, N., and Izadi, S. 2009. Putting the physical into the digital: issues in designing hybrid interactive surfaces. In *Proc. of BCS-HCI '09*.
12. Kirsh, D. 1995. The intelligent use of space. *Artificial Intelligence*, v.73 n.1-2 (February), 31-68.
13. Kirsh, D. 2009. Problem solving and situated cognition. In P. *Robbins and M. Aydede*, editors*, Handbook of Situated Cognition*, 264-306.
14. Kirsh, D. 2009. Projection, Problem Space and Anchoring. In *N. A. Taatgen & H. van Rijn (Eds.)*, In *Proc. of the 31st Annual Conference of the Cognitive Science Society*, 2310-2315. Austin.
15. Kirsh, D. and Maglio, P.P. 1994. On distinguishing epistemic from pragmatic action. *Cognitive Science* 18, 513-49.
16. Lee, J., Post, R., and Ishii, H. 2011. ZeroN: mid-air tangible interaction enabled by computer controlled magnetic levitation. In Proc. of *UIST '11*. ACM, NY, USA, 327-336.
17. Marshall, P. 2007. Do tangible interfaces enhance learning? In *Proc. of TEI '07*. ACM, NY,USA, 163-170.
18. Marshall, P., Cheng, P.C.-H., and Luckin, R. 2010. Tangibles in the balance: a discovery learning task with physical or graphical materials. In *Proc. of TEI '10*. ACM, NY, USA, 153-160.
19. Mazalek, A. and Hoven, E., van den. 2009. Framing tangible interaction frameworks. *Artif. Intell. Eng. Des. Anal*. Manuf. 23, 3 (August), 225-235.
20. Norman, D.A. 2010. Natural user interfaces are not natural. *interactions* 17, 3 (May 2010), 6-10.
21. Patten, J. and Ishii, H. 2000. A comparison of spatial organization strategies in graphical and tangible user interfaces. In *Proc. of DARE '00*. ACM, NY, USA, 41-50.
22. Piper, B., Ratti, C. and Ishii, H. 2002. Illuminating clay: a 3-D tangible interface for landscape analysis. In *Proc. of CHI '02*. ACM, NY, USA, 355-362.
23. Shaer, O. and Hornecker, E. Tangible User Interfaces: Past, Present, and Future Directions. *Foundations and Trends in Human-Computer Interaction*, v.3, 1-137.
24. Soute, I., Kaptein, M. and Markopoulos, P. 2009. Evaluating outdoor play for children: virtual vs. tangible game objects in pervasive games. In *Proc. of IDC '09*. ACM, NY, USA, 250-253.
25. Vallée-Tourangeaua, F. and Penneya, A.K. 2005. The impact of external representation in a rule discovery task. *European Journal of Cognitive Psy.*, 17.
26. Underkoffler, J. and Ishii, H. 1999. Urp: a luminous-tangible workbench for urban planning and design. In *Proc. of CHI '99*. ACM, NY, USA, 386-393.
27. Weller, A., Villejoubert, G. and Vallee-Tourangeau, F. 2011. Interactive insight problem solving. *Thinking & Reasoning*, 17(4):424-439.
28. Wikipedia Authors. Connect Four. (Four-in-a-row). In Wikipedia. Retrieved November 21, 2012, from http://en.wikipedia.org/wiki/Connect_Four
29. Xie, L., Antle, A.N., and Motamedi, N. 2008. Are tangibles more fun?: comparing children's enjoyment and engagement using physical, graphical and tangible user interfaces. In *Proc. of TEI '08*. ACM, NY, USA, 191-198.
30. Zaman, B., Vanden Abeele, V., Markopoulos, P. and Marshall, P. 2012. Editorial: the evolving field of tangible interaction for children: the challenge of empirical validation. *Personal Ubiquitous Comput*. 16, 4 (April), 367-378.
31. Zhang, J. 1997. The nature of external representations in problem solving. *Cognitive Science*, 21, 179-217.