Thinking with the Body
David Kirsh  (kirsh@ucsd.edu)
Dept of Cognitive Science
University of California, San Diego

Abstract
To explore the question of physical thinking – using the body as an instrument of cognition – we collected extensive video and interview data on the creative process of a noted choreographer and his company as they made a new dance. A striking case of physical thinking is found in the phenomenon of marking. Marking refers to dancing a phrase in a less than complete manner. Dancers mark to save energy. But they also mark to explore the tempo of a phrase, or its movement sequence, or the intention behind it. Because of its representational nature, marking can serve as a vehicle for encoding aspects of a target movement, marking as a method or priming neural systems involved in the target movement, and marking as a method for improving the precision of mentally projecting aspects of the target.

Keywords: Marking; multimodality; thinking, embodied cognition, ethnography.

1. Introduction
This paper explores how dancers and choreographers use their bodies to think about dance phrases. My specific focus is a technique called ‘marking’. Marking refers to dancing a phrase in a less than complete manner. See fig. 1 for an example of hand marking, a form that is far smaller than the more typical method of marking that involves modeling a phrase with the whole body. Marking is part of the practice of dance, pervasive in all phases of creation, practice, rehearsal, and reflection. Virtually all English speaking dancers know the term, though few, if any, scholarly articles exist that describe the process or give instructions on how to do it.¹

When dancers mark a phrase, they use their body’s movement and form as a representational vehicle. They do not recreate the full dance phrase they normally perform; instead, they create a simplified or abstracted version – a model. Dancers mark to save energy, to avoid strenuous movement such as jumps, and sometimes to review or explore specific aspects of a phrase, such as tempo, movement sequence, or underlying intention, without the mental complexity involved in creating the phrase ‘full-out’.

Marking is not the only way dancers ‘mentally’ explore phrases. Many imagine themselves performing a phrase. Some of the professional dancers we studied reported visualizing their phrase in bed before going to sleep, others reporting mentally reviewing their phrases while traveling on the tube on their way home. Our evidence suggests that marking, however, gives more insight than mental rehearsal: by physically executing a synoptic version of the whole phrase – by creating a simplified version externally – dancers are able to understand the shape, dynamics, emotion, and spatial elements of a phrase better than through imagination alone. They use marking as an anchor and vehicle for thought. It is this idea – that a body in motion can serve as an anchor and vehicle of thought – that is explored in this paper.

It is a highly general claim. It has been said that gesture can facilitate thought, [Golden Meadow 05]; that physically simulating a process can help a thinker understand a process [Collins et al 91], and that mental rehearsal is improved by overt physical movement. [Coffman 90] Why? What extra can physical action or physical structure offer to imagination? The answer, I suggest, is that creating an external structure connected to a thought – whether that external structure be a gesture, dance form, or linguistic structure – is part of an interactive strategy of bootstrapping thought by providing an anchor for mental projection. [Hutchins, 05, Kirsh 09, 10]. Marking a phrase provides the scaffold to mentally project more detailed structure than could otherwise be held in mind. It is part of an interactive strategy for augmenting cognition. By marking, dancers harness their bodies to drive thought deeper than through mental simulation and unaided thinking alone.

Hand Marking

Fig 1a
In Fig 1a an Irish river dancer is caught in mid move. In 1b, the same move is marked using just the hands.

Hand Marking

Fig 1b
River dancing is a type of step dancing where the arms are keep still. Typically, river dancers mark steps and positions using one hand for the movement and the other for the floor. Most marking involves modeling phrases with the whole body, and not just the hands.

¹ Search by professional librarians of dance in the UK and US has yet to turn up scholarly articles on the practice of marking.
2. Methodology

To explore the role of physical activity in dance cognition we were fortunate to study the creation of a new dance piece by the noted choreographer Wayne McGregor, the resident choreographer of the Royal Ballet in London. WM created the dance we studied with his own company, Random Dance, a group of ten extremely talented dancers. An eleventh dancer from a different company in Europe joined the group for the first period of dance creation.

The dance company’s process of creation occurred in two phases: a three week episode at the University of California, San Diego (UCSD) in the winter of 2009; and a second period in London, in the late summer of 2009, just preceding the official première at Sadler’s Wells Theater.

**Method:** During each phase, written notes were taken in real-time. During the UCSD phase, fifteen students took notes; during the London phase, a single experienced ethnographer took notes. Both phases, UCSD & London, were exhaustively videotaped using five high definition video cameras placed on the walls, and, whenever possible, two standard video cameras were placed on the ceiling. The whole rehearsing process, 11AM to 5PM, five to six days a week was captured. Video footage exceeds 110 hours (times 5-6 cameras) and captures all scheduled interactions between choreographer and dancers during the dance making process.

Cognitive ethnography requires acquiring a detailed knowledge of a community of practice, and then using that knowledge to illuminate specific episodes of activity. [Williams 06]. To acquire knowledge of the community of practice we interviewed the choreographer as well as the dancers repeatedly. We also reviewed all notebooks, and used our interviews as an opportunity to discuss specific moments of creative activity. The choreographer was interviewed for between forty and sixty minutes on digital video each morning and night. The dancers were interviewed at the end of each rehearsal. Our aim with the dancers was to have them reflect on specific elements of the rehearsal that day, and wherever possible, to show us through movement the dancerly decisions they made. Four dancers were selected and interviewed for thirty minutes each day. About 70 hours of interviews, in total, were videotaped.

To code the video we used ELAN, a free software system developed by the Max Planck Institute for Psycholinguistics, designed originally for studying gesture and small-scale interactions. Systematic audiovisual analysis depends on having a well-defined vocabulary of coding – a classification of activity and phenomena. After a few days of ad hoc coding a formal vocabulary was established by the whole team (20 people) to characterize ongoing activity. After the UCSD phase of capture, we reviewed the video data and selected special phenomena, such as marking for more detailed coding. In the London phase, we interviewed dancers explicitly about marking to probe them on their own views about marking. These interviews were undertaken in addition to the normal 30 minute ones we conducted. In several such sessions, we had the dancers come before the camera and dance *in full* a phrase they knew well; we then asked them to show us several ways they might *mark* that same phrase, and to describe the reasons they would mark one way versus another. We also interviewed them in a less structured manner, often returning to the question: “When do you mark, and how?” which led to multiple follow up questions and nuances of speech, as well as spontaneous performances from the dancers. The videotaped answers, with the corresponding gestures and markings, were transcribed and analyzed in detail with ELAN. On this basis, we created a hierarchical taxonomy of marking, yielding the three parent groups reported below. Inter-coder reliability in distinguishing these parent marking types exceeded .9, on a sample of 25 video snippets of marking among our most experienced coders (n=3).

3. The Gross Function and Structure of Marking

At the highest level, three *functions* of marking can be distinguished.

1. **Marking-*for-self*:** dancers use their body to encode an aspect of a phrase for themselves. This may be for reinforcing memory, reflecting on sequence, or for scrutiny of spatial relations, among other reasons.

2. **Marking-*for-others*:** dancers use their bodies to encode an aspect of a phrase that others can focus attention on. For example, before a new performance, choreographer, choreographic assistant, and lighting manager review all phrases on stage for space.

3. **Joint-marking:** two or more dancers run through a phrase as a tightly coupled team, verifying timing and grips jointly for each other.

**Small vs. Large Marking**

![Fig 2a](image1.png) ![Fig 2b](image2.png) ![Fig 2c](image3.png)

Figs 2a, 2b, 2c show the contrast between small and large marking. In 2a, a male dancer is remembering a step, using his hand to small mark it. In 2b, a female dancer is showing how she marks a pirouette. She uses a formal gesture for a pirouette that she learned as a ballet dancer. Her marking is small and conventional. In 2c, a second female dancer marks a phrase using movements that are of comparable size to those in the full phrase. She is clearly modeling the phrase.
There are also a few things to note, at the highest level, about the *structure* of marking.

**Variability of size:** Marking comes in a continuum of sizes, from very small to full size (but less energetically). In ‘small marking’, the amount of movement is minimal; the marking movements tend to be in the upper body (hands and head mainly), and the objective is to review the steps, the relationship between simultaneous movements (arm and leg together), and occasionally to attend to timing. See figs 2a and 2b. In extreme cases, such as Irish river dancing (fig 1), marking may be done exclusively with two fingers marking foot rhythm, position, and movement. When marking is very small, it is a form of gesture. In larger marking, especially when the function is to show the floor space required by a movement, or to show off the structure of a phrase to someone else, the movements may be full size but with less intent, emotion, or energy than the real movement (fig 2c). They are imperfect models of the complete phrase, but lacking certain attributes, such as intensity, motion dynamics, or fine detail.

**Substitutability:** A movement in one body part can represent the movement in another. Hand movements and head tilts regularly stand for the motion of different body parts: a hand movement may represent a leg movement, a head turn may represent a torso turn or a whole body turn; if the legs perform in parallel, one leg may stand in for two. This too is shown in figs 2a and 2b. See figs 3a, b for a standing version and fig 1 for finger version.

**Idiosyncratic vs. Conventional Marking**

![Fig 3](image3.png)

In 3a a dancer marks a leg movement with his hands in his own idiosyncratic manner that is a hybrid of conventional ballet marking and personal style. In 3b A dancer from a strong ballet tradition offers a conventional small marking with her hands.

**Conventional:** In classical ballet and other formalized dance forms, dancers are taught to use specific gestures as ways of marking certain moves. These are a conventionalized form of small markings. For instance, as seen in fig 2b, the female dancer marks for the interviewer with her hand to show that, at a certain point in the phrase, a pirouette is required. In fig 3b she shows us a gesture for a *pas de bourrée*. These small gestures refer to a complex sequence of full moves well known by ballet dancers. We observed that dancers who do not rely on a ballet vocabulary still mark in a way that is reminiscent of ballet marking: but each dancer has personal idiosyncrasies that violate convention. In fig 3a, for instance, a dancer with deep training in both modern and ballet represents a leg movement with his arms, a hybrid marking that is part conventional and part personal gesture.

**Aspectival:** Marking typically represents an aspect of the full phrase, with some forms of marking focusing solely on tempo, others focusing on sequence, still others focusing on spatial position. For instance, when dancers mark for space they will keep the scale of the full phrase, but other aspects will be ignored or only partially represented, such as the dynamics of the phrase. At other times, just the movement of the upper body or the torso orientation may be marked and the movement of a leg or arm is left completely unmarked. Evidently, when dancers mark they are attending to only certain aspects of the phrase.

4. Analysis

Is it plausible to see marking as a vehicle of thought? There are a few promising ways to approach this question. Perhaps the most obvious line is that marking is a type of gestural semiotic system, possibly like a linguistic code. If gesture can function as a vehicle of thought, as some have argued, then why not marking?

It is useful to classify gestures according to where they lie on ‘Kendon’s Continuum’ (McNeill 92). At one extreme, there are “gestures of the kind that Kendon has called ‘quotable’ … gestures that must be configured according to pre-established standards of form in order for them to function as signs, such as the OK sign among North Americans” (McNeill & Duncan 2000). These are compositional and behave in many respects like words or phrases in a language. At the other extreme are ‘gesticulations’. These are idiosyncratic, created on the fly, and motivated by imagery rather than convention.

In dance, marking in the classical tradition of ballet is convention-driven and quotable. Despite individual differences in marking style, dancers still conform to general norms. Although marking conventions vary from ballet company to company, it does not take long for a professional dancer to pick up the idiosyncrasies of a company. This suggests there are rules determining the structure of ballet marking, and that local differences in marking style should be viewed as akin to differences in accent or handwriting. They need to be learned but are not different in principle than dialects of a common language.

In contemporary dance, the *reference* of marking – the phrases full-out, or aspects of those phrases – are not easily segmented. Movements in contemporary dance are freer, often novel. There are also far fewer conventions governing how dancers should mark. But not none. In the group we studied, for instance, there were quite strict rules about how to mark for the choreographer or his assistant. The spatial
dimensions of the phrase were to be preserved, though energy, and pace could be lessened.

The implication is that marking might well lie nearer the language side on Kendon’s continuum than the gestural side. This needn’t be a surprise. If there are written notation systems for encoding dance, such as Laban notation, then as long as marking is as expressive as these notation systems, anything that can be encoded on paper can be encoded through marking. The one requirement is that there be semantic rules for interpreting the notation and semantic rules for interpreting marking.

It is here, however, that the analogy with language fails. Marking is a reliable language only when (a) the dancers are marking for others – the other forms of marking lack adequate semantic rules; and, (b) only when the point of marking is to display space, position, and structural form, all aspects of the full-out phrase that the choreographer or his assistant can directly see in the marking itself. If the point of marking were to call attention to movement sequence or to motor preparation, external observers would often be unable to infer the movements being sequenced or prepared for.

This is perhaps the key point. If someone states, “there is a circle with radius 30 meters”, a competent interpreter need not have seen such a circle beforehand to know what the sentence means. It is enough to know the meaning of the terms ‘circle’, ‘radius’, ‘30 meters’ to generate an interpretation. That is what semantic rules are for. By contrast, in marking, because there is so much idiosyncrasy in marking when dancers are marking for themselves, or when marking an aspect of a phrase that is not visibly similar to the full-out phrase (space), observers cannot ‘see’ the full-out move ‘in’ a marked version unless they already know what the full-out looks like. This explains why dancers rarely, if ever, mark a phrase they do not already know, and why choreographers never request dancers to show them novel phrases by marking – they insist on a full-out. Evidently, both parties need a clear idea of the target in advance of the marking. They have to have seen the full-out phrase to be able to ‘project’ it from its marking.

I believe this proves that much if not the majority of marking is not language like. It relies on prior acquaintance with the target, and then matching the mark to its target. That process more closely resembles a pattern completion process than a generative process of constructing the target. Languages are essentially generative, the point of marking is to avoid generating the whole target.

But if marking does not behave as a language this raises a paradox: if a dancer, or an observer, needs a clear idea of the full-out phrase in order to correctly interpret its marked version, why bother with the marking? How can marking ever be more powerful than inner visualization or imagination alone? What more can the physical manifestation of a movement add to the target already ‘mentally grasped’ through imagination?

One answer is that physical movement is helpful when one wants to measure the distance covered in a phrase. External distance is not guaranteed to be accurate in a mental representation. [Ledermen 87]. And there may be other physical dimensions available in the physical execution of a phrase that are only implicit in its mental representation (for instance, the physical tension in leaping off the floor or lifting another person).

But, beyond making physical attributes measurable, [see Kirsh 10], what extra cognitive benefits can physical marking provide that surpass mental rehearsal?

Here are two possibilities. They offer a different take on how marking might serve as a vehicle of thought.

1. Marking is a way of anchoring projection to a target. By providing a marked version of a target, a dancer can project a better representation of the target than imagination unaided. Marking, therefore, is a causally important way of augmenting thought. It is a component of a distributed vehicle of thought, consisting of an inner part and an outer part, which enables clearer thoughts. (cf. Hutchins 05)

2. Marking is a way of priming the neural system of a dancer, thereby enhancing imagination (or projection) by activating cortical elements that would be involved in the full-out movement. Marking is a way of enhancing the vividness and detail of imagination.

Marking as a method of anchoring projection. In the phenomenology of perception, a distinction can be drawn between perception, projection, and imagination. See fig 4.

- When we perceive an object, our experience is that we are seeing an object that is really there; we feel it is what causes our perception.
- When we project onto an object, we experience ourselves intentionally augmenting the object; we feel we partially cause our experience.
- When we imagine an object, we feel as if we are the sole cause of our imagined experience.

Fig 4. The difference between perception, projection, and imagination is represented here by three conditions of a tic-tac-toe game. Perception: subjects see moves. Projection: subjects see only the tic-tac-toe grid, and mentally augment it with moves. Imagination: subjects see a blank page and all aspects of the game are imagined – no external stimuli to scaffold or structure imagination.

The application to marking is shown in Fig 5. If the full-out phrase is represented by the complete triangle in 5a, marked versions are represented by 5b – 5e. The marked versions are either fractions or distortions of fractions of the full. But they support projection to full-out, if one has been exposed to the full-out already.
This form of projection is not a standard completion process. In completion, the target is a superset of the fragment. For example, tang___ is a stem that supports completions like tangent. The fragment ta_g_s supports the completions targets or tangles. In both cases, the target completes the fragment. In projection, the structure that augments the fragment need not complete it because it may produce a new structure that has none of the subset structure. For instance, in 5c, the completion is larger in all dimensions except corner angle. In 5d and 5e, even the angles are not preserved. Projection is not completion.

Kirsh [09] showed that it is easier to conceptualize a target, or recover more memory of a target’s structure, if there is something outside that one can ‘lean on’ for support. It is easier to project than to imagine if there is something helpful outside to support the projection. Recall is better for projected imagery than imagined imagery [ibid].

Marking as Projection

![Diagram of marking as projection](image)

Fig 5. The idea of marking as a sequence of illustrations of decreasing verisimilitude to the full phrase. 5a: a complete path at full scale. 5b: same path, full scale, shown by vertices and directions. 5c: smaller path, the interpreter must now know the scaling function. 5d: a stylized version of 5a. 5e, a smaller version of 5d, interpreter must project both shape, angles, and know the scaling function.

The relevance to marking is that when dancers mark, they may be creating a physical scaffold that facilitates projection. This would explain what ‘extra’ a dancer gets by physically marking a phrase rather than mentally rehearsing it. They get an external structure they can extrapolate from. This enables them to generate a conception of the final target that is more vivid, complete, and requiring less mental effort, than when they mentally rehearse without the support of overt movement. Moreover, dancers are able to choose how much extra memory support they want, just by marking more completely. When their mental image of the target is already clear, their marking may be minimal. When they have a weak mental image of the target, they may mark it more extensively, thereby increasing the vividness and control over their conception of the target.

Marking as a method of priming. A second benefit of marking may be that it involves more brain activity than mental rehearsal alone. It may facilitate muscle memory of details or deeper processing of movement goals.

The importance of muscle memory in dance is part of standard teaching. Muscle memory refers to the system of motor procedures – motor schemata – that have been stabilized through practice and are activated during performance. [Krakauer 06] Initial movements prime later movements. Priming also facilitates projection. Priming refers to an increased sensitivity to a stimulus due to prior exposure to a related stimulus. For instance, subjects who recently hear, see, think, and especially perform a particular movement will recognize aspects of that movement, sooner than those who have not. [Koch et al 04] The extent of priming is also a function of the depth of processing involved in the earlier exposure. [Challis, 92, Smith et al 83]. A person who thinks hard about a dance phrase – its energy, sequence, rhythm or spatial extent – will prime more choreographic relatives of the phrase, and prime them more deeply, than someone who merely sees the phrase briefly. Since motor preparation, spatial planning, and proprioceptive monitoring are involved in marking, it is likely that even more areas of cortex are involved in marking than in mental rehearsal alone. This suggests that during marking, there will be more opportunities for deeper processing – more chance to see deeper relations among movement components – than during mental rehearsal. Marking should prime the phrase more deeply, making it easier to remember it in the future.

If marking helps a dancer to envision the target phrase better, it helps to explain why marking is beneficial. Given the importance of internal processes, however, marking is best understood as the external part of an internal-external process. It is best seen as the external part of a distributed vehicle of thought.

5. Conclusion

I have argued that marking is a form of physical thinking. A dancer creates a partial version of a phrase, attends to it while creating it, and because of processes like priming and projection, the dancer is able to understand something deeper about the phrase’s structure than through imagination alone. When dancers mark, they are closely coupled with the dance product they are externalizing. They rely on that product to think with. Their performance of the marked phrase is part of their ongoing process of grasping the phrase. In some ways, their relation to marked material is reminiscent of what E. M. Forster (27) said about language: “How can I know what I’m thinking until I see what I say”. For Forster, the external vehicle of a thought – its linguistic formulation – was a real time achievement of putting the thought into words. It made the thought more precise in virtue of the constraints of language. There was no point asking whether the articulated content was the same as some internal version already encoded in an internal language intrinsically understood, as suggested by Fodor (75) and others. For Forster, as well as for Wittgenstein (51), the articulation is part of the thinking process.

My suggestion, here, is that for a dancer, Forster’s rhetorical question can be rephrased as: “How can I know
what my phrase really is until I see what I do?” A dancer’s thought of his or her phrase is partly shaped by what is marked. Dancers do think about their phrases without dancing them or marking them. But, by marking-for-self dancers think better about their full-out phrase. Physical movement replaces mental computation. Instead of imagining transformations, they execute them externally. Marking is part of a distributed vehicle of thought with internal and external parts closely coupled.

Acknowledgements: I gratefully acknowledge help from Dr. Dafne Muntanyola on ethnographic analyses of the dance data, from twenty five students in my class on Creative Cognition in Dance, from Wayne McGregor and the dancers in Random Dance, from Scott delaHunta the director of Random Research, and from the Committee on Academic Research, UCSD for a seed grant for expenses.

References
Projection, Problem Space and Anchoring
David Kirsh (kirsh@ucsd.edu)
Dept of Cognitive Science, UCSD
La Jolla, CA 92093-0515

Abstract

When people make sense of situations, illustrations, instructions and problems they do more than just think with their heads. They gesture, talk, point, annotate, make notes and so on. What extra do they get from interacting with their environment in this way? To study this fundamental problem, I looked at how people project structure onto geometric drawings, visual proofs, and games like tic tac toe. Two experiments were run to learn more about projection. Projection is a special capacity, similar to perception, but less tied to what is in the environment. Projection, unlike pure imagery, requires external structure to anchor it, but it adds ‘mental’ structure to the external scene much like an augmented reality system adds structure to an outside scene. A person projects when they look at a chessboard and can see where a knight may be moved. Because of the cognitive costs of sustaining and extending projection, humans make some of their projections real. They create structure externally. They move the piece, they talk, point, note, represent. Much of our interactivity during sense making and problem solving involves a cycle of projecting then creating structure.

Keywords: Projection, interactivity, imagination, sense making, cost structure, externalization, visual thinking, situated cognition.

Introduction

Why do people typically perform better by staring at a chessboard, a tic tac toe board or a geometric proof and project what they might do, rather than memorize the board or proof as it is initially, then close their eyes while they think of possibilities? When subjects consider possible moves in a chess game, one popular account is that they are searching a problem space; they are exploring a purely mental representation of the game’s states, entertaining possible actions and evaluating consequences. This way of speaking leaves unexplained the relation between the physical board that is perceived and the mental process of searching an internal representation. The two might be uncoupled. And in fact, masters rarely need the cognitive support provided by a physical chessboard. They can do all the work in their heads. So a purely mental representation seems apt for them. But less expert players do benefit from a board’s presence. They interactively coordinate their projections – their simulation of what if’s – with the board as they see it outside. Why does a board help them project? How?

My real concern here is with interactivity: how, when and why do people interact with their environment when making sense of situations, solving problems and so on. I present a truncated account of what I believe is a key, perhaps the key interactive process in reasoning and sense making: the projection-create-project cycle. I believe this cycle lies at the heart of much sense making, especially problem oriented sense making. It lies, as well, at the heart of most planning and tangible reasoning. A complete analysis of these phenomena would require the simultaneous study of behavior and brain. My analysis here is confined to the fine grain of behavior, involving scrutiny of the details of what people do when they make sense and reason.

In videographic studies of people understanding such things as illustrations, instructions, models and diagrams we found that subjects typically find ways of interacting with at-hand tools and resources – often in creative ways – to help them make sense of those targets. Sometimes these sense-making actions are as simple as gesturing or pointing with hand, body or instrument, muttering while looking, marking or note taking, or shifting the orientation of the target. Sometimes they involve talking with others. When tools are placed near subjects – manipulable things such as rulers, pencils, and physical parts of models – we found subjects regularly use these as ‘things to think with’. They use them to create or supplement local structure to facilitate projection and mental experimentation. This is the heart of the project-create-project cycle: use what is perceived to help you do what you can in your head – namely, try to understand things by projecting possibilities, by somehow augmenting what you see – then externalize part of that mentally projected augmentation so that you free up cognitive resources. This process of externalization simultaneously changes the stimulus and makes it easier to project even deeper. If tools make it easier to externalize what you are thinking, then tools are used. This cycle of projecting, externalizing, then projecting again continues as long as subjects stay focused – though as with any exploratory or epistemic process a subject may soon loop, get stuck, or run out of novel projections. Let me define some terms and properties.

Projection: The basic idea

Projection is a way of ‘seeing’ something extra in the thing present. It is a way of augmenting the observed thing, of projecting onto it. In contrast to perception, which is concerned with seeing what is present, projection is concerned with seeing what is not present but might be. It is sensitive to what is present yet sufficiently controlled by a subject to go beyond what is perceived.

In figure 1 two rather different illustrations are displayed. The first – a cartoon – requires subjects to interpret the symbolic meaning of the key elements. The image must be recast as a ‘keyframe’ in a narrative invented by the reader, in this case, a narrative of retirees watching helplessly as
their pension money is lost forever to inflation. If you ask subjects to tell you what precedes this image and what is likely to follow they usually offer a brief story, as if running a movie forward or backward. The account typically involves a few cartoon frames describing events leading up to the current situation, but more importantly it usually has a gloss about the meaning of the loss, which is not highly visual. In observing narrative projection we found that people rarely have an urge to mark up the picture.

In figure 1b we see a geometric illustration. It is meant to show the givens for the question: will the line extending from A through D bisect BC? To solve the problem subjects scan the figure to interpret the labeling and invariably interpret the claim by imaginatively projecting a line from A through D and into BC. They may estimate magnitudes, such as whether the projected line cuts BC into equal lengths, or whether BD splits angle ABC in two. At some point, though, they are likely to reach for their pencil to add constructions or annotations to make their mental projections physical. They mark angles as equal, segments as bisected, lines as parallel or perpendicular, triangles as congruent. This interactive process of projecting and marking continues until they solve the problem or they clutter the figure so badly that they cannot keep in mind what goes with what. At that point they erase dead ends and backtrack to an earlier point in their search process.

Both illustrations 1a and 1b involve projection of meaning, but in 1b subjects also seem to be augmenting what they see. I am concerned in this paper with this second type of sense making, where subjects project quasi-perceptual structure – imagery of sorts – onto the target as if marking it up. The first example, framing and embedding the image in a narrative, is a worthy sense-making topic; but my focus is on projection that is more perception-like. It is a form of projection that typically leads to physically creating new structure.

**Figure 1.** Two illustrations: 1a is a narrative illustration requiring the viewer to make a sensible story out of the image. It involves identifying narrative worthy elements and interpreting them. For instance, the birds are not part of the narrative content but the lost money is. 1b is a figure showing a few geometric constructions. In one type of math problem, subjects are given a linguistic statement of a problem; they convert the key premises into visible shapes, and then using the figure as an aid they prove certain truths, such as that a line through AD will bisect BC. The diagram is used to clarify the givens and support inferences and allowable augmentations (new constructions). Both phases of diagramming – the conversion of linguistic to visual form, and the construction of additional lines and property labels – involve projection. In the first phase, before a subject inks the figure in the first place s(he) usually formulates a partial plan concerning where to draw the lines and how they will look. In the next phase, conjectures are often tested by projection, as if ‘seen’ through augmented reality, before pen is once again taken to paper.

**Relation to Perception and Imagination**

Projection, perception, and imagination lie on a continuum of stimulus dependence, with perception being the most dependent and imagination the least.

**Perception** is strongly dependent on the physical stimulus it is about. We cannot see what is not there. Even on those occasions where we have a perceptual experience of something that is, in fact, not there, such as the illusory edges shown in the first portion of figure 2, the experience is justified by the stimulus. Sometimes, real objects do produce that very effect. For instance, a solid white triangle occluding a black edged triangle would create the illusory edges shown. So the presence of perceptual mistakes and illusions is consistent with perception being stimulus dependent. Our perceptual system has been designed to recover real structure. It is tightly coupled to the outside.

**Projection** is also dependent on present stimuli but much less so than perception. The coupling is looser because projection offers a peak into the possible, into what could be there, or what might be useful if it were there, but is not. It is like wearing augmented reality glasses. But with one difference. In staring at a chessboard and seeing how a knight might move, a subject must mentally remove the knight from its current spot. The layout must be changed as well as augmented.

![Image](image_url)
It is an open empirical question how much external structure must be present and for how long it must persist for it to serve as the substrate of projection. But there has to be something present to project onto, something to anchor projection. (cf. Hutchins 2005). To grant the mind the flexibility it constantly reveals, the anchoring structure need not be persistent – we can anchor a thought in a person’s gesture or in the direction a person points. But whether persistent or ephemeral there must be some external structure present, else there is nothing to distinguish projection from pure imagery. Whether that structure is enduring enough to provide a stable understructure to support repeated projection – as a chessboard does for our projection of possible moves – or whether it has a fleeting presence, triggering a single projection but then is gone, the stimulus enabling the projection in both cases has a reality outside the agent.

What can be said about the coupling between projection and substrate, between projection and anchor? First, projections are most often momentary. Some may persist in mind for minutes but usually they do not. Visual markers nicely demonstrate this temporal distribution. A visual marker or FINST, as Pylyshyn (1994) calls them, is the internal counterpart of a physical indicator laid down to mark an object or location. FINST’s help us keep track of multiple objects in the visual scene. If two bees, similar enough to be perceptually indistinguishable, were buzzing around our den, and we wanted to keep an eye on each, our visual system would conjure up two FINSTs to mark the bees, thus providing the extra structure needed to keep the two distinct. FINST’s of this sort have enough mental persistence to provide the stability needed for referential thought. “I wonder if that bee is going to land beside the other?” But once the tracking task ends the FINST’s are released. And more often than not the need for simultaneously tracking two objects in one’s nearby visual environment is not long lived. So if FINST’s are representative projections they do not endure long.

In chess, the duration of projection too may vary, though the structure projected is more complex than a FINST. Most often, if a novice relies on a board to facilitate projection the projected move is considered briefly, then rejected. The ones that are not summarily dismissed form the first step in a chain of moves, so they must persist at least as long as the time it takes to create the chain. In both cases, despite their differences, projection relies on external structure being present. This contrasts with chess masters who do not need the board to think about moves. They operate more in a ‘virtual reality’ of their own making rather than the ‘augmented reality’ I am introducing as the mark of projection. And their board is conceptualized and chunked to a much greater degree.

*Imagination* better describes the chess master’s mental activity. Their representation of the board and current situation is completely sustained internally and they have control over what they imagine next. Imagination is often defined as “a mental representation of a nonpresent object or event” Solso (1991 p.267). In psychological accounts of imagery (e.g. see Denis, 1991; Kosslyn, 2005) mental images have two primary dimensions: vividness and controllability. Vividness refers to the clarity, "sharpness" or sensory richness of an image (Richardson, 1999). Controllability refers to the ease and accuracy with which an image can be mentally transformed or manipulated (Kosslyn, 1990). We may assume that masters have both vivid and well-controlled images of chess situations.

**Externalization: part of the project-create cycle**

Externalization is a way of taking information or mental structure generated by an agent and transforming it into epistemically useful structure in the environment. It is a way of materializing structure that first was mental – it is the create part of the project-create-project cycle.

Externalizations are everywhere: annotations, notes, constructions in geometry, gestures, utterances, encoding order in layout, (Kirsh 1995, 2008) etc. Often the action of externalizing alters the information or projection in useful ways. This is a key factor in thinking with things, in knowing what you are thinking by seeing what you are saying, and so on. Externalizations may leave persistent traces, as in annotations or rearrangements, or they may be present only during the externalization process, as when someone gestures or talks while thinking.

Externalizations always serve an epistemic function. But they also may have pragmatic consequences too. A chess move is at once an externalization of an inner projection and a move in the game. And of course there are other actions that change the environment in epistemically useful ways that are not externalizations: registration of maps, turning on the news channel, etc. These are actions that alter the epistemic landscape of activity but they do not bear the right relation to internal activity to qualify as externalization, and they are not part of the project-create-project cycle.

**When does external structure help performance?**

Externalizing a mental projection allows a person to release at least some of their working memory, replacing it with perceived structure. So it serves as an effective interactive strategy for increasing mental power. The value of this interactive method is easy to appreciate when the structure being created is something like a construction in a geometric proof: A construction typically starts out first as a mental projection and then, if it seems fruitful, is materialized by marking the illustration. But when is structure necessary to improve performance? Some chess masters can play equally well with eyes closed. When does externalizing help?

For example, would staring at a blank tic tac toe board while calling out moves help performance? A blank board does not carry any state information. How could it help?

To answer that question we ran a few simple experiments, video’ed and analyzed performance to see how behavior and cognitive strategy differs when a board is present from when it is not. We used a 3 by 3 tic tac toe board first, then we scaled the game to a 4 by 4 board to see if the complexity of
the game affected the value of external structure. Our conjecture was that having an empty tic tac toe table would help in both cases.

![tic tac toe table](image)

**Figure 4.** Training image and the three conditions.

**Experiment One. Procedure:** In figure 4 the training stimulus and all experimental conditions are shown. Experimenter and subject took turns calling out numbers 1 to 9 corresponding to the cells in a 3 by 3 board. Subjects could not mark their paper but could gesture if they wished. The goal was to get three in a row defined in the classical tic tac toe manner. Subjects were given an initial training period during which they mastered the translation of number to position on the table. A within subject design was used. Each subject played in each of the three conditions: blank, table, and table + XO. There were three games to a condition, three conditions to a block and two blocks to an experiment. All conditions were counterbalanced within and between subjects to control for order effects and microgenetic learning. In the table condition, subjects were given a sheet with a blank tic tac toe table to view if they so wished. In table + XO a similar tic tac toe table was given to subjects but with the letters X and O above it. In the blank condition subjects were given a blank piece of paper to look at. This was meant to serve as the imagination case – the unanchored problem space case. During our pilot study with 7 subjects we found that performance varied considerably among subjects. In particular, there were a few subjects who regularly did best on the blank condition. During debriefing it was apparent they had good imagery abilities. So all pilot subjects were called back and given a standard imagery test: vividness of visual imagery questionnaire (VVIQ-2) by D. Marks (1995). In the end 27 subjects were run and all tested with VVIQ-2.

**Results.** As shown in Table 1 the mean time to make a move was relatively close in all conditions and statistical tests showed no significant differences between conditions. Apparently, seeing a table does not help in 3 by 3 tic tac toe. When we divided the subject pool into strong visualizers – the upper 33% of our VVIQ scores - and weak visualizers – the bottom 33%, there were differences in means but none that were statistically significant. We also checked for order effects, to see if subjects showed significant learning during the experiment. None was noted. Nor were their significant order effects (microgenetic learning) among strong and weak visualizers. Other individual differences were more suggestive, however. Fully half of our subjects actually did better on the blank condition than the table condition. This was significant (p=.002). This difference does not correlate with visualization ability.

**Discussion.** Coming into the 3 by 3 case, and on the basis of our pilot data, we assumed that staring at an empty tic tac toe table would help subjects - at least weak imagers - because we thought an empty table would function as an aid to memory. Without a table an agent must remember the structure of the table as well as the values in all its cells. So having a table to observe ought to reduce memory load.

Apparently, our conjecture is wrong in the case of 3 by 3 tic tac toe. Overall, nothing is to be gained by projection. Imagination is just as good. Either the memory task is not challenging enough to warrant offloading memory, as it is in chess where the board and piece configuration contains a huge amount of information, or subjects are already at ceiling.

There is, however, another possibility. Projection is an expensive process. It requires anchoring imagined elements – mental X’s and O’s – with physical locations. There is no a priori reason why mental tic tac toe elements should easily fit the physical table subjects look at. Some might like a large table others a small one. Indeed, several subjects reported a disconnect between their imagery, or their mental imagery strategy, and the table they were asked to use. One reason there is no general effect in 3 by 3 tic tac toe, then, is that, for many subjects, the benefits of projecting may not override the costs. For those subjects, projection is not a good strategy.

Some support for this interpretation can be found from the surprising finding that a full 50% of our subjects actually did worse in the table condition than in the imagery condition. What might explain that other than posting a cost to projecting – a cost to anchoring? See table 2.

**Table 1.** Mean performance 3 by 3. Shorter is better. Differences are not significant.

<table>
<thead>
<tr>
<th></th>
<th>All Subjects</th>
<th>3 by 3 Means:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Weak Visualizer</td>
<td>Strong Visualizer</td>
</tr>
<tr>
<td>Blank</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Table</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Table + XO</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2.** Many subjects found it easier to play the 3 by 3 game in their imagination.
To test whether there is a threshold where the urge to use external structure to support projection becomes overwhelming, we ran a second experiment in which we scaled up tic tac toe to a 4 by 4 table. Intuitively everyone has a visualization limit. Once that limit is reached the cost of projection is more than paid back by the reward in memory saving, or visualization reliability, or reduced mental effort.

**Experiment Two.** Procedure: cells are identified with numbers 1-16 and learned beforehand. The goal of this enlarged game of tic tac toe is to be the first to get four in a row.

To illustrate better the importance of having a table to work with becomes more valuable, especially for weak visualizers.

**Results:** Unlike the 3 by 3, all subjects in the 4 by 4 reported that external supports were helpful. But their empirical performance did not always confirm this claim.

**Discussion:** These findings are consistent with the hypothesis that projection has a cost that is offset once the complexity of a problem passes a threshold. The differences we observed between weak and strong visualizers suggest, further, that this threshold varies considerably with imagery capacity. Of course this does not prove that projection happens whenever the going gets rough. It suggests that how necessary it is to project depends on both the information size of a problem, and the ‘effort’ an individual must expend in anchoring imaginary elements.

Several qualitative observations add to this picture. During debriefing interviews several subjects reported that they used different strategies in the table and blank conditions. When no table was present they felt overwhelmed and played defensively, using a strategy of blocking the opponent as quickly as possible rather than trying to win. Imagery alone is hard. In the table condition, however, strong 3 by 3 players – those who typically are better visualizers – initially believed they could project enough state to play offensively. They felt that with the support given them by a table they could compete with an experimenter who played with paper and pencil. Invariably subjects made errors and soon shifted their table strategy to a defensive one, and their table performance improved, soon becoming their best condition. What is interesting is that they believed in their projective ability and that they could endure the mental effort of following a harder strategy than they would consider in the pure imagery condition. When this strategy proved unreliable they fell back on using the table with a defensive strategy for greater reliability and speed. (Fewer errors were made in the table condition but not significantly so).

A second qualitative observation we made concerns the number and type of gestures made in 4 by 4 versus 3 by 3 games. It soon became apparent that the more difficulty a subject had with the tic tac toe task, even with the help of a table, the more likely they were to externalize state information to help them out, in this case with hands and fingers. Humans are ingenious at finding ways of overcoming internal state limitations. They invent methods of reducing the overall cost of performing a task, especially when the alternative is failure. They project then create structure.

For example, subject M, found a clever way of placing his fingers on the cells and the lines between the cells in the 4 by 4 table to encode more than 10 cells worth of information. Obviously he would have had far more difficulty encoding this information without the table there to ‘lean’ on since he would have had to project a visual structure with lines and cells ‘under’ his fingers.

There is much more to be said here concerning the nature of coding with hands and gesture and the timing of these interactions. But it may be more worthwhile tying this study back to the question of how people use projection to make sense of diagrams such as visual proofs and illustrations of mechanical systems.

The idea I am exploring, is that projection is related to perception, perhaps continuous with it, though it cannot be identical with perception because it is directed at augmenting the world. You cannot see what is not there. Yet in some theories of perception, most notably enactive theories, O’Regan and Noe (2001), Noe (2004) perception already contains a component of ‘seeing the future’. For instance, when we see an object we do not literally see it, though it cannot be understood in this special way, because it is a process of increasing the priming level of some of the things we would see if we were
to act in certain ways. The relevance to tic tac toe and to
to geometry is that because we are able to label a tic tac toe
table by writing on it, we have a weakly primed version
of the labeled table already in mind. The stronger our
disposition to add labels the stronger is the priming for
seeing a table augmented with those labels. Projection is a
way of intentionally increasing the level of primed states
of the world. It lets us anticipate what the world would be like if
we did act to make it so.

Now consider the visual proof shown in figure 6a that the
sum of $\frac{1}{2^n}$ converges to 1. As you immerse yourself in the
proof do you feel you are recreating a progression of cuts?
Do you see that the operation of halving a square whose
sides are 1 by 1, and then halving the remainder (whether it
be a square or a rectangle) is a recursive process that will
never yield a structure larger than the original square? I
contend that this quasi-simulation of cutting is a form of
projection and lies at the heart of making sense of visual
proofs.

![Figure 6](image)

**Figure 6.** Two different types of visual reasoning
are at play in 6a and 6b. 6a shows a geometric proof
and requires understanding the recursiveness implicit
in cutting regions in half. 6b requires physical
understanding of the effect of pushing down on a
lever.

Now look at figure 6b. Is more force or less force
required to lift the load when the fulcrum is moved closer to
the foot? How did you find the answer? By mentally
moving the fulcrum and then simulating the consequences
for the foot? This projected animation cannot be perception
because, presumably, perception requires that what you see
is, in fact, there to be seen. You cannot see the future. Yet
there is something perception-like in this projected imagery
even if it is not nearly as vivid as perception. Again
projection seems to lie at the heart of our sense making
visual thinking here. As we found with tic tac toe, our
proxy for thinking in abstract problem spaces, projection
lets us probe problems by tying our thinking to external
structure. It lets us anticipate how the world might be,
when we act on it.

**Conclusion**

I have been arguing that projection is a basic cognitive
capacity involved in visual thinking, in much problem
solving and in making sense of illustrations, diagrams, and
many types of planning and reasoning situations. Projection differs from imagination and imagery in being
anchored to visible structure. When we project it is like
wearing augmented reality glasses: we lay structure over
existing structure. There are no doubt other modalities of
projection beside vision, but I have not considered these
here. An experiment was presented in which subjects were
tested to see if they performed better when there was more
structure of the right sort to anchor their projections. As
predicted, when useful anchoring structure is present
subjects score more highly. The usefulness of such anchors
depends on a subject’s imagery ability and the complexity
of the problem. For tasks that are simple relative to a
subject’s imagery ability, external anchors are of no value.
Projection is replaced by imagination. But as a task
increases in complexity, projection and anchoring becomes
important even for good imagers. At some point everyone
benefits from external structure. The costs associated with
projecting and anchoring are offset by the returns derived
from mental ease, memory saving and reliability.

The relevance of this to problem solving and our opening
question concerning chess players should be obvious:
masters who can play chess in their imagination do not need
to look at a board when playing. They have so overlearned
the chess board and possible configurations that they can
play equally well with or without a board. But less
practiced players need a chessboard and pieces. They cannot
sustain a meaningful problem space for chess without the
help of perception to provide anchors to a real board. Their
problem space is more a projection than autonomous mental
space. If anchors do not exist they must create them.

**References**

Wheatsheaf.


Kirsh, D. *The Intelligent Use of Space*. Artificial Intelligence,
Vol. 73, Number 1-2, pp. 31-68, (1995).

of Situated Cognition*. New York: Cambridge
University Press. 264-306.

Neuropsychology (22) 333-347.

*Components of high-level vision: a cognitive neuroscience
analysis and accounts of neurological syndromes*. Cognition.
34(3):203-77.

research". *Journal of Mental Imagery*, 19, 153-167.

O'Regan, J. Kevin and Noe, Alva (2001) *A sensorimotor
account of vision and visual consciousness*. Behavioral and
Brain Sciences. (24) 939-973


Hove, U.K.

Balan