The Tetris model of the information seeking process

Max L. Wilson

Department of Computer Science
Swansea University, UK
m.l.wilson@swansea.ac.uk

Abstract. Although many attempts have been made to model the information seeking process, so that we might better understand it and improve search systems, previous models have typically tried to flatten the dynamic and repetitive sequence of actions into a single set of stages. Some have surpassed the linear models with circular connections to show that users may, for example, search, examine results, and then refine their search, and repeat. Here we argue that simply adding circular attributes to the model does not fully capture the variability in the search process. Instead, we present the Tetris model of the information seeking process, as a means to more completely capture both the activities and the dynamic process involved in searching. After presenting the model, and related work, we continue by describing the additional benefits provided and how it may better inform design of information seeking systems.

1. Introduction

Understanding the Information Seeking Process (ISP) remains a hard challenge, despite the decades of research that has been produced by the Information Seeking and Retrieval community. In fact, libraries have been modeling and hypothesizing about seeking behaviour well before the invention of the computer, in order to enable visitors in finding the books they desire [1]. While many popular ISP models have been generated (e.g. [2-4]), nearly all have been supplied with a caveat: that users may jump back and forth through the process at will. This common caveat represents a common flaw in such models: that users do not progress through linear phases, but jump between several active, passive, and reflective states in an unpredictable fashion. This largely unpredictable switching is seen clearly in the ISP presented by Marchionini [4], shown in **Fig. 1**. In fact, the multitude of arrows highlight that the linear progression from left to right only represents the ideal or best path that can taken by a searcher and emphasizes that users are rarely able to take this best path.

In this article, we present a Tetris model of the information seeking process, where we agree with the states that are widely accepted among the various models, but not necessarily that they are a sequence of sequential stages. Instead, the temporal progression captured in the Tetris ISP model, allows for the reality that users transition between these states in almost any order. Here, the process of start to finish is controlled by the resolve of information needs, rather than having progressed through a set of stages. Simply, the different stages, or what we consider to be states, are represented as Tetris layouts with varying complexity, where more information must be fitted together to reach their goal. The larger information need, motivating the IS process, is complete when the Tetris board has been cleared of pieces.

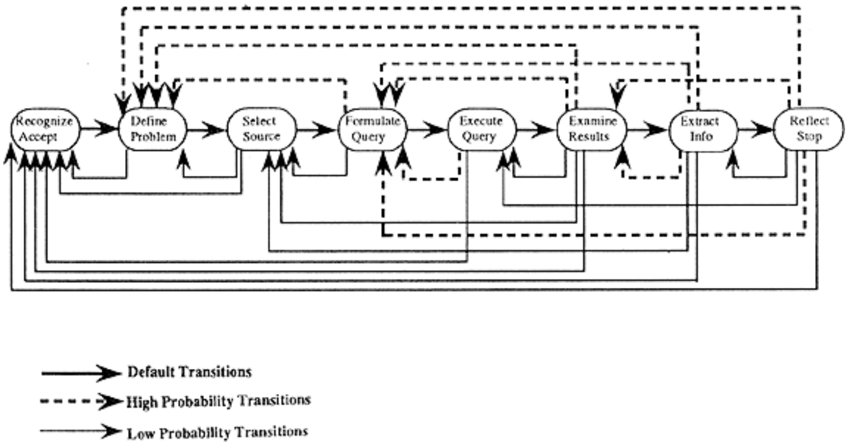

Fig. 1. The Information Seeking Process provided by Marchionini [4]

In the following sections we present related work on Information Seeking (IS), models of IS, and their weaknesses. We follow the discussion of related work with a description of the Tetris model and how it responds to the weaknesses of the other models. Finally, we conclude with a discussion of its benefits, uses, and potential future developments.

2. Related Work

Information seeking is widely accepted as the process of resolving a specific information need. Belkin et al. describe this information need as an Anomalous State of Knowledge in their ASK model [5]. Dervin et al. [6] represent this need as a gap over which the user has to build a bridge, assuming that they want to get across it. Such higher level models, usually considered as general Information Behaviour models, assume that a user has alternatives to information seeking, where they choose to ignore a gap or even reject the gap [7]. This Tetris ISP model is not a model of general information behaviour, and assumes that a user chooses to resolve their information need.

Another brand of model, such as that proposed by Saracevic [8] and Järvelin and Ingwersen [9], capture the levels of context and influence that surround information seeking activities. In Saracevic's model, search activities are influenced by their knowledge, situation and environment. Similarly, the model shown in **Fig. 2** by Järvelin and Ingwersen, presents similar levels of context that can each affect the search process. Notably seeking, which encompasses basic information retrieval, is affected by both the nature of the driving work task, as well as their working, social, and cultural contexts. Again, this Tetris model does not aim to capture levels of context like these models, but focuses on the seeking context and the processes taken within it. Notably, though, we again see many bi-directional arrows in **Fig. 2** indicating that there is no one fixed process taken by users.

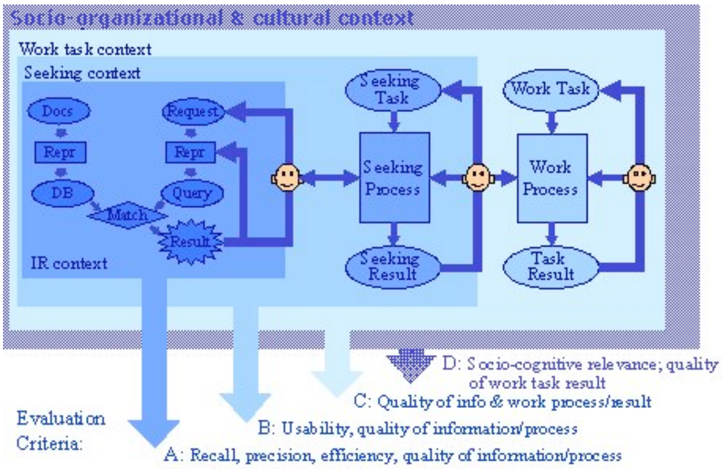

Fig. 2. Searching behaviour is made up of multiple layered contexts, where simple information retrieval is the most narrowly focused; figure from Järvelin and Ingwersen [9].

Many models of the ISP itself have been proposed and largely agree on the default phases or stages that people progress through. **Fig. 3** shows a comparison of two key process models by Ellis [2] and Kuhlthau [3], visualised together by Wilson [10]. By constructing this figure, Wilson noted the similarity of the two process models, which, incidentally, follow a similar process to Marchionini (shown in **Fig. 1**). Essentially, they each share the notion that searchers start with a realisation point, try to define their problem, perform some searching or browsing actions, analyse the results they receive, and stop when their need has been resolved. Wilson discusses these and other information seeking models in more detail.

While most ISP models do attempt to represent the search process as a progression through linear steps, attempts have been made to generate non-linear models. Foster presented one such non-linear model, which is described as analogous to an artists palette, where similar stages to the other models are freely available to the searcher to use as needed. As a searcher/artist, defining (and re-defining), formulating, searching, analysing, and reflecting are used as needed to finish the job.

Spink's [11] model of information seeking, shown in **Fig. 4**, clearly shows that the search process is made up of iterative cycles that contain feedback loops of searching, retrieving, and judging of results. Feedback loops and cycles appear regularly in models to cater for the random and non-linear activities in search, including the sensemaking [12] and information foraging models [13].

While most of the key ISP models awkwardly cater for, ignore, or even abstract-out the fact that users switch frequently between stages, Marchionini's [4] model is the most explicit in representing the reasons and conditions in which changes occur. Marchionini "crudely", as he describes it, models these switches by identifying both more and less likely paths that users may follow backward through the stages. Further, the absence of arrows between certain states implicitly highlights switches that do not occur. Another example to explicitly consider state changes was provided by Belkin et al [14], who generated detailed 'scripts' for 16 different types of searchers, and identifying script entry and exit points where people can switch between searcher types. The full descriptions of these search episodes and scripts are extensive, and also fairly rigid, despite allowing users to transfer between them.

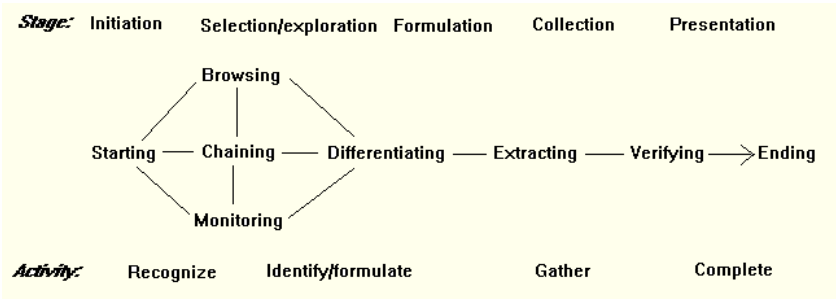

Fig. 3. A comparison of Ellis' (center flow chart) and Kuhlthau's (top sequence) search stages, as visualised together by Wilson [10].

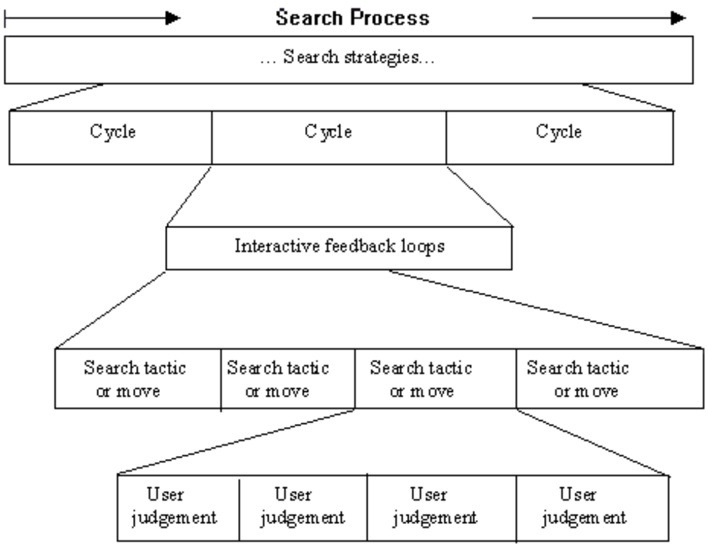

Fig. 4. Spink's model of the search process

The Tetris model, described below, is built upon this assumption by modelling progression, time, and state-changes independently, as opposed to the typical left-to-right models described above which typically confound all three into one.

3. The Tetris Model of the Information Seeking process

Here we present a model of the ISP that uses the game of Tetris as an analogy for the types and sequences of actions involved in search. The aim of the game, shown in **Fig. 5** is to fit the descending pieces, of varying shape, together so that they create one or more complete lines across the width of the screen. When a complete line has been created, the line is removed from the game and the score is increased. In order to better fit these pieces together, users are able to rotate the descending objects and move them left and right as necessary. To make this game fun and challenging, the rate at which pieces descend increases, so that the user has less time to a) work out where to place the piece, and b) move and orientate the piece accordingly.

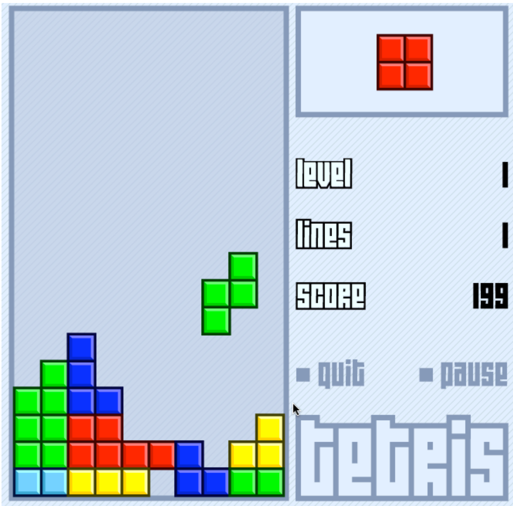

Fig. 5. A game of Tetris

We are not the first to study Tetris for academic purposes. Kirsh and Maglio [15], for example have studied the difference between epistemic and pragmatic actions performed by users, as the speed of the game increases, to learn more about perception and reaction protocols. Further, Veksler and Gray used a Tetris-based task set to measure learning [16].

As an overview of how we use this Tetris analogy, the game window itself is considered to be a working space for an information need. Descending pieces in the game are considered as new pieces of information arriving into working space. The pieces already at the bottom of the game window, making up incomplete horizontal lines, represent a current information need. Any completed lines of the game are removed, and a score value incremented, which represents the cumulative amount of knowledge a searcher has about this topic. The depth of the incomplete horizontal lines represents the depth of the current information need.

While time, in this Tetris analogy, is simply modelled as linear, progression is modelled separately as the completion and removal of lines in the game window. The ideal situation is that a user is able to resolve an information need by completing and removing all lines that are currently visible.

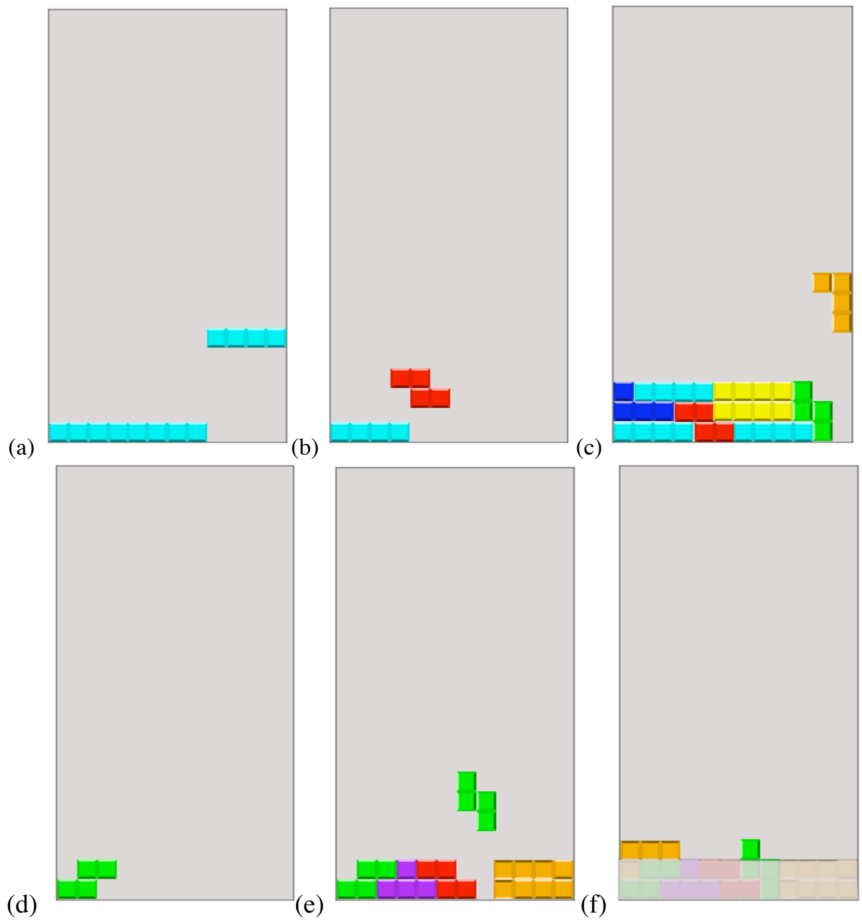

Fig. 6. Example information seeking states, where (a) is a simple lookup task resolved easily; (b) is a learning example where the information need was deepend by the second bit of information; (c) presents a deep information need being eventually solved with one piece; (d) shows a more complex initial information need for investigation; (e) shows an investigation need getting even deeper; and (f) represents excess information found that is surplus to solving the information need.

The different stages that searches may be in are identified by the state of the game. If a new piece of information arrives into an empty window, then a user has a new information need to resolve (typically the first stage of every ISP model). The user thus begins a new information seeking episode as they try to clear that new piece of information from the board. As the user analyses search results, new pieces arrive in the window of the game, and hopefully help to resolve the information need. As the user can see what information need they have, they can construct an idea of what remaining pieces they might need. As the pieces arrive, the user has to analyse how and where they fit into their developing knowledge on the topic. In the game of Tetris, rotating and moving the pieces from side to side achieve this. These new pieces, however, may not directly resolve the need, and deepen the number of lines that need resolving, causing the searcher to reformulate their information needs. As the pieces do fit together, however, the user is able to reflect on the state of the game. When the lines have been cleared, the user can consider that they have resolved their information need.

In the following subsections, we further explain this analogy with three diverse examples that represent different complexities of information needs, taken from the broad categories described by Marchionini [17]: Lookup, Learn, and Investigate.

3.1 Lookup example

A simple lookup has been shown to be a regular part of everyday life, when a simple problem is identified (caused by the arrival of new piece of information), described easily, found quickly, and therefore resolved with ease. This scenario, portrayed in **Fig. 6**.a, can be represented easily by the first three pieces of a game each being made up of 4 horizontal blocks. Each block can be laid side by side, horizontally, so that they make up an entire line and the problem is removed. That is, the original problem, the first piece, was simple and, given the empty game window, created a shallow information need. Equally, the searcher came across the appropriate information to resolve the need easily. The newly found information fits simply, without further deepening the information need, and the line (or information need) is resolved. Overall the IS episode was quickly resolved and the screen, until a new problem piece arrives, is entirely empty.

3.2 Learning example

In learning examples, an original information need may appear to be relatively simple, such as needing to buy a camera. The next piece found, however, indicates that more is to be learned before a final camera can be chosen, such as the difference between metering modes and ISO capabilities. The new pieces of information confound the chance of resolving the problem in a single line, and deepen the problem space to a number of Tetris lines (**Fig. 6**.b). Consequently, more information must be sought to resolve any or, hopefully, all of the lines. It is possible, however, for this deepened information need to be completely solved with a single final piece, by clearing several lines at once (**Fig. 6**.c). Alternatively, each of the lines may be resolved one at a time with successive pieces.

3.3 Investigating example

Investigating is represented by tasks such as planning, analysis, synthesis, and evaluation, where the complexity lies in the initial problem pieces, such as a multifaceted problem, or a poorly defined information need. If the initial problem piece has a complex shape (**Fig. 6**.d), then the information need can only be resolved by several pieces. The depth of the problem, unlike Learning examples, is controlled by the complexity of the earlier rather than the later pieces. This does not exclude, however, the later pieces from increasing the complexity of the problem, as investigating may often include learning (**Fig. 6**.e).

3.4 Life long learning example

As a final example of the versatility of modelling the ISP with a Tetris analogy, life long learning can be modelled by the game of Tetris. There may be some topics that people spend their entire lives learning about, such as the focus of an academic career, or a personal interest or hobby. During the large period of time, people may engage in lookups, learning exercises and investigations, where each one resolves something they did not understand before, or discovers new pieces of information. Realistically, we process many bits of information as we search, and by resolving one information need, there may always be unresolved blocks that represent information that has not been explained or investigated (**Fig. 6**.f). Here, we can imagine that users pause the game until such time that they wish to engage in information seeking to either resolve new information needs, or the leftover blocks from a previous information seeking session.

In Tetris, users' unresolved lines might be occluded by new pieces of information (e.g. **Fig. 5** above). The searcher may have to clear several lines and resolve several information seeking problems before one day resolving something that they did not complete before. This may occur when new pieces of information need to be checked, or first understood before the original problem can be resolved.

4. Discussion

In the following section, we discuss the proposed Tetris model under three topics. First, we compare the proposed Tetris model against previous models of IS, so that we can understand the new contributions it provides. In light of these contributions, we discuss how the model can help researchers or designers to potentially improve their work. Finally, we discuss whether Tetris has any reverse analogies back into life.

4.1 Comparison to other models of the search process

There is a general consensus within research that the ISP involves a start goal, whether it be broad or focused, and an endpoint where all or part of the goal has been achieved. This may occur within one IS episode [14], or across multiple search episodes [18, 19]. In the Tetris model, the start and completion of an ISP is not represented by an axis or a direction. Instead, a problem starts when a new piece enters an empty screen, part of a problem is resolved when a line is completed and it is removed from the game window, and a whole problem is completed when the screen is made clear again. Further, by using the depth of the unresolved lines to represent the depth of a problem, we are able to easily model that any bit of new information has the ability to either resolve or further deepen the problem. Finally, the model does not require any particular order of need identification, description or information discovery, as each can occur at any time to be involved in the game of an overall seeking process.

4.2 Using the Tetris model in system design

As said above, one aim of the Tetris model is to capture the well-observed behaviour that users regularly repeat any or all of the stages involved in information seeking and that stages as a term, in fact is misleading. Users will regularly experience evolutions in the way they understand their problem, define it, express it, and resolve it. Consequently, the thesis of the Tetris model is that we should not be modelling incremental stages but the fact that each of them has the potential to *deepen* the ISP, and that completion of it (if at all) occurs as the full depth of the problem is resolved.

This concept of depth, rather than incremental stages, is a fairly novel way of thinking about the ISP, and so may immediately have consequences for new designs of search systems. For example, the model poses the question of whether, instead of tracking the stage at which a user is currently at, a system could track the level of depth and detail that the user is currently at. There is very little need, for example, to provide a growing overview of a domain, if the user simply performs a short look-up. A space for synthesising information, however, would be very useful as the users information need deepens. Marchionini and White [20] report on previous research that has shown, for example, that automatically trying to create a synthesised view of a domain allowed users to perform Learning and Investigative tasks significantly faster, but provided no significant benefit for look-up tasks. Further, a faceted browser called RB07 [21] assumes by default that users will keyword search (a simple look-up) and then refine results with facets of metadata if they need to explore the results in more detail. Google appears to share this sentiment, by typically placing recommended query refinements after the first ten results. some interfaces exist that support different depths of search, however they have typically been designed from the progressive models. We have yet to see, to the authors knowledge, a system designed specifically from the view of depth, rather than progression through stages.

Pragmatically, we hope that academics will, as we will, review their search systems whilst considering how different example Tetris games would be supported. Such analysis could lead to improvements in supporting users seamlessly transitioning from what was perhaps expected to be a quick look-up to more exploratory forms of search.

4.3 Taking the Tetris analogy further

In contrast to the life-long learning example (Section 3.4), there are topics that individuals have learned about in the past, but have many blocks still to resolve. If the user is not actively seeking, one way or another, in those topic areas, then these games can be considered on pause. Un-pausing may, therefore, occur as new information arrives passively, or if the searcher actively engages in a new search episode.

One open research question for the Tetris model is to decide if users are able to fail games of ISP-Tetris, as they reach maxima in their understanding of topics. In Tetris, failing is indicated by the stack of unresolved, incomplete, lines filling the available vertical space allocated to the game. Perhaps individuals who fail to understand hard problems have reached their vertical limit on games, despite having resolved some set of lines in the past. Renewed effort, however, can often surpass current limitations of understanding, and so perhaps users are starting a new game on the same topic, building with a fresh screen but with a history of the lines they resolved last time. Another unanswered question revolves around users finding information they already know. In the Tetris model, we suggest that known information does not arrive as pieces in the game window.

In the example of life-long learning, we cannot be sure whether searchers have successfully kept their Tetris game relatively clear, or simply have applied more effort and so have increased the available vertical space. If the latter is true, then we may see that determined learning or investigation in a topic might be supported as extending the vertical space of a game and trying to resolve some of the problem lines. Those that give up on topics, however, choose not to increase their vertical space, but are content with the history of lines they have resolved, and ignore the amount of lines that they have not. In these terms, and in the notion of pausing discussed briefly above, we can consider that this Tetris analogy can model some aspects of general information behaviour.

One final element of the Tetris game that is not covered by this model, is that of the increasing speed at which pieces descend as time progresses. This element of speed is added to the game to introduce both challenge and enjoyment. In real life information seeking is often performed under time pressures such as deadlines or medical emergencies. Similarly, many user studies impose an element of time on their information seeking tasks. Users are often timed as they carry out the tasks, and most analyses assume that improvements in task completion time make their system better. In some respects this is true, if the system in question is producing pieces that will fit together nicely to resolve a problem, then they have supported the users by not accidentally deepening the information need. If the original challenge is more exploratory or investigative, however, then time pressure may make it harder for searchers to fit new pieces in with the unresolved lines. In the real world, searches often need time to reflect on what they have found and to decide if it relates to what they have discovered already.

Going against the norm for information seeking user studies, several recent studies have used increased search time during exploratory tasks as a positive measure, as it means that users have potentially resolved more aspects, of lines, and of their information need [22, 23]. We can only consider increased search time as a positive measure, however, if the quality of their task output, such as a report on a topic, has also improved. In the majority of user studies, the answer and its potential quality are fixed, and so reduced time is then considered as positive for achieving a known goal. From this discussion of time, it is clear that the Tetris model once again can help to describe information seeking behaviour, explaining both why reduced time and increased time can be used as positive measures, depending on the nature of the task.

5. Conclusions

In this article, we have proposed a Tetris model that describes the process of Information Seeking. Unlike most models of Information Seeking, which usually try to reduce the process to a single progression across stages, the Tetris model describes a progression separately from time, whilst allowing users to freely repeat any stages of: need identification, formulation, expression and information discovery, as required. By making the depth of the problem analogous to the current depth of the unresolved lines in the game, the model allows for any new information in the process to either resolve or deepen the problem. Further, by making the resolving of part or all of a problem analogous to completing a whole line across a Tetris board, the start and end of an ISP can be identified by the exit or entry to an empty screen, respectively.

In our examples of lookup, learning, and investigating, we have shown the versatility of the Tetris model for describing diverse IS scenarios. Further our discussion has identified its potential strengths, as well as how these strengths might be used to support the design of future search interfaces. Finally, we discuss the potential for the model to have overflowing analogies to areas such as increased user effort and failure.

As identified in the introduction, the purpose of any model is to help us structure what we do know and hypothesise about users, in this case engaging in the ISP, so that we can identify new avenues of research. In turn these new avenues of research can support or strengthen our models, through identifying and resolving limitations, so that they may again inform new ideas. Consequently, we intend to analyse the findings of our own research and the research of others into user studies and noted IS behaviour. By carrying out such analyses, we can either find support for the Tetris model of search, or identify limitations for further investigation. Ultimately, it is likely that we will discover both evidence for and against the model, as all models and hypotheses are abstractions of the truth, and so by its very definition do not cover every aspect of real human behaviour. Regardless, we hope the dialogue of doing so will continue to increase the field's collective understanding of what is involved in the Information Seeking Process.

6. References

1. Dewey, M., *A classification and subject index for cataloguing and arranging th ebooks and pamphlets of a library*. 1876: Fascimile reprinted by Forest Press Division, Lake Placid Education Foundation in 1976.
2. Ellis, D., *A behavioral model for information retrieval system design*. Journal of Information Science, 1989. **15**(4-5): p. 237-247.
3. Kuhlthau, C.C., *Inside the search process: Information seeking from the user's perspective*. Journal of the American Society for Information Science, 1991. **42**(5): p. 361-371.

4. Marchionini, G., *Information Seeking in Electronic Environments*. 1995: Cambridge University Press.
5. Belkin, N.J., *Anomalous states of knowledge as a basis for information retrieval*. Canadian Journal of Information Science, 1980. **5**: p. 133-143.
6. Dervin, B., L. Foreman-Wernet, and E. Lauterbach, *Sense-Making Methology Reader: Selected Writings of Brenda Dervin*. 2003: Hampton Press. 400.
7. Godbold, N., *Beyond information seeking: towards a general model of information behaviour*. Information Research, 2006. **11**(4): p. Paper 269.
8. Saracevic, T. *Modeling interactio in information retrieval (IR): a review and proposal*. in *Proceedings of the American Society of Information Science*. 1996.
9. Järvelin, K. and P. Ingwersen, *Information seeking research needs extension towards tasks and technology*. Information Research, 2004. **10**(1): p. paper 212.
10. Wilson, T.D., *Models in information behaviour research*. Journal of Documentation, 1999. **55**(3): p. 249-270.
11. Spink, A., *Study of interactive feedback during mediated information retrieval*. Journal of the American Society for Information Science, 1997. **48**(5): p. 382-394.
12. Russell, D., et al. *The cost structure of sensemaking*. in *Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems*. 1993. Amsterdam, The Netherlands: ACM Press.
13. Pirolli, P. and S. Card. *Information foraging in information access environments*. in *Proceedings of the SIGCHI conference on Human factors in computing systems*. 1995. Denver, CO, USA: ACM Press.
14. Belkin, N.J., et al., *Cases, scripts, and information-seeking strategies: on the design of interactive information retrieval systems*. Expert Systems with Applications, 1995. **9**(3): p. 379-395.
15. Kirsh, D. and P. Maglio, *On distinguishing epistemic from pragmatic action*. Cognitive Science, 1994. **18**(4): p. 513-549.
16. Veksler, V.D. and W.D. Gray. *State definition in the Tetris task: designing a hybrid model of cognition*. in *Proceedings of the Sixth International Conference on Cognitive Modelling*. 2004. Pittsburgh, PA, USA.
17. Marchionini, G., *Exploratory search: from finding to understanding*. Communications of the ACM, 2006. **49**(4): p. 41-46.
18. Mackay, B. and C. Watters. *Exploring multi-session web tasks*. in *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*. 2008. Florence, Italy: ACM Press.
19. Morris, D., M.R. Morris, and G. Venolia. *SearchBar: a search-centric web history for task resumption and information re-finding*. in *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*. 2008. Florence, Italy: ACM Press.
20. Marchionini, G. and R.W. White, *Find What You Need, Understand What You Find*. International Journal of Human-Computer Interaction, 2007. **23**(3): p. 205-237.
21. Capra, R. and G. Marchionini. *The relation browser tool for faceted exploratory search*. in *Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries (Demonstration)*. 2008. Pittsburgh, PA, USA: ACM Press.
22. Capra, R., et al. *Effects of structure and interaction style on distinct search tasks*. in *Proceedings of the ACM/IEEE-CS Joint Conference on Digital libraries*. 2007. Vancouver, British Columbia, Canada: ACM Press.
23. Kammerer, Y., et al. *Signpost from the masses: learning effecs in an exploratory social tag search browser*. in *Proceedings of the 27th international conference on Human factors in computing systems*. 2009. Boston, MA: ACM Press.