Applying Formal Design Methods to Serious Game Design: a Case Study

Ivo Bril[1], Nick Degens[1]
[1]User-Centred Design, School of Communication, Media & IT, Hanze University of Applied Sciences, Groningen, the Netherlands
bril.ivo@gmail.com
d.m.degens@pl.hanze.nl

Abstract: An important step in the design of an effective educational game is the formulation of the to-be-achieved learning goals. The learning goals help shape the content and the flow of the entire game, i.e. they provide the basis for choosing the game's core (learning) mechanics. A mistake in the formulation of the learning goals or the resulting choice in game mechanics can have large consequences, as the game may not lead to the intended effects.
At the moment, there are many different methods for determining the learning goals; they may be derived by a domain expert, based on large collections of real-life data, or, alternatively, not be based on anything in particular. Methods for determining the right game mechanics range from rigid taxonomies, loose brainstorming sessions, to, again, not any method in particular. We believe that for the field of educational game design to mature, there is a need for a more uniform approach to establishing the learning goals and translating them into relevant and effective game activities.
This paper explores two existing, non-game design specific, methods to help determine learning goals and the subsequent core mechanics: the first is through a Cognitive Task Analysis (CTA), which can be used to analyse and formalize the problem and the knowledge, skills, attitudes that it is comprised of, and the second is through the Four Components Instructional Design (4C-ID), which can be used to determine how the task should be integrated into an educational learning environment. Our goal is to see whether these two methods provide the uniform approach we need. This paper gives an overview of our experiences with these methods and provides guidelines for other researchers on how these methods could be used in the educational game design process.

1. Introduction

It is becoming more and more apparent that serious games are here to stay. Clark et al.'s (2015) analysis of the current state of the field shows us that games *can* be effective, but require more in-depth research into what makes them effective. At the moment, the body of empirical evidence for the effectiveness of games in terms of learning outcomes is too small to be able to provide the answer to this 'what' question (Bellotti et al., 2013). This has led to a lack of standardized procedures in the design, development, and assessment in serious game design research and hampers the maturation of this field as a whole (Mayer et al., 2014). The multitude of possible design approaches available to researchers right now and the lack of details about how to use these approaches causes researchers to use them in heterogeneous ways (Mayer et al., 2014).

One way of dealing with this problem would be to change *how* we as researchers discuss our design in (academic) writing. Gaydos (2015) discusses this in his work and makes a stand for more elaborate discussions on how we deal with complications that occur throughout the design process. The main reason for the importance of this elaboration is to make the design process more insightful for the reader and to reflect more on the design choices and their impact on the end result. Doing so will possibly prevent a repetition of previously made errors and will strengthen the underlying literature on *why* (and when) an educational game can be effective. A good example of what such a discussion would look like is Clark et al.'s (2016) summary of a research project about teaching Newtonian physics using educational games. While their paper is perhaps quite unique as it discusses almost a decade of research, design, and development, they discuss their mistakes and problems (e.g. clinging on to a popular game mechanic) and how they dealt with them.

Another way of dealing with the heterogeneity problem would be to introduce more formal approaches into the field and to see whether that formalisation can be used to develop a more universal design-approach. One such formal approach has gained some traction in recent years; the four component instructional design (4C-

ID) model of Merriënboer, Jelsma, and Paas (1992). Two games have been designed using this empirically-validated design model:

- Van Rosmalen et al.'s (2014) CHERMUG. This collection of smaller games aims to teach the user how to do quantitative and qualitative research and the aspects that play a role in research in general (e.g. statistics); and
- Van Bussel et al.'s (2014) CRAFT. A game focused on teaching technical knowledge and skills in vocational education (mechanic mechatronics).

Both games used the instructional design model to guide their game's design, providing an interesting idea: why try to create new, game-specific, frameworks when older, validated, models can provide the same result?

A different existing method that has also gained some traction is cognitive task analysis (CTA). This method which is used often in designing lessons for tasks which involve complex decision making was also used for the CHERMUG game (Boyle et al., 2012).

As we have discussed in earlier work, we wish to find ways in which to properly formalize the relationships between three aspects: the user, the game mechanics, and learning goals (Degens, Bril, and Braad, 2015). In line with our view, we would like to provide our insights on using formal methods to establish a solid basis for the game design based on the *interaction* between these three aspects.

In this paper, we will discuss this ambition from the perspective of a case study: the development of a serious game aimed to support health care professionals by improving their physical behaviour throughout their job's tasks. In Chapter 2, we will introduce the case in more detail. This is followed by a description of the Cognitive Task Analysis that was executed to better formalize the domain knowledge in Chapter 3. In Chapter 4, we will describe the use of the 4C-ID model to better structure the game design flow. We conclude the paper in Chapter 5 with a discussion about the intricacies of using these non-game methods for serious game design ends, and provide some guidelines for future researchers to apply these methods.

2. The Case

Bronkhorst et al. (2014) report that nurses in the Netherlands suffer both in a physical and mental manner, due to their workload, working posture, and other factors. The physical and mental strain they endure during their work causes nurses to call in sick for a longer period of time or, even worse, to experience burn-out symptoms.

Looking after patients is a very physically straining task which requires one to be vigilant of one's own physical health. This includes properly executing the actions necessary to take care of a patient, minimalizing the strain on the caretaker him- or herself. Bronkhorst et al.'s (2014) report states that 85% of the nurses experience physical pain or other physical problems, which would imply that they do not work safely from an ergonomic perspective.

To help caretakers identify and deal with these strains, a group of developers, researchers and domain experts has started working on the development of an e-learning tool (codenamed LudoVita). The goal of this project is to provide caretakers with the education and training they need to deal with the physical and mental strain they encounter during their work. The tool itself will be comprised of short serious games, each focussing on helping the caretaker learn how to deal with a specific aspect that causes the mental or physical strain, and other types of interventions. This paper will discuss the problem of physical strain due to wrong posture, and how serious games could play a role in preventing chronic injuries.

The first step to understand the specific problems that the nurses encounter is by analysing their daily routines and behaviours, formulating specific learning goals from that analysis, and then using the learning goals to properly structure the serious game. As the tasks themselves were documented quite clearly, the only things left to analyse were the decisions and thought processes of the caretakers themselves. For this, we used a Cognitive Task Analysis.

3. Cognitive Task Analysis

When designing an intervention (i.e. training) for a complex task, knowing what underlying decisions and lines of thinking play a role in executing the task is quintessential. Cognitive Task Analysis (CTA) is a method comprised of a multitude of ways to analyse complex tasks using subject matter experts (SME). The resulting analysis provides a base from which a training method can be designed; a good overview of the state of CTA and what it can be used for is provided by Clark et al. (2008).

Educating the nurses on the proper execution of their tasks required us to understand the reasoning and justification behind each task; when to do the task, why, and how. As the Bronkhorst et al.'s report had shown that nurses were not working safely, we needed to first verify if this was due to a lack of understanding or if there were other, yet unknown, reasons.

3.1 Method for CTA

3.1.1 Design

The CTA was designed to gain a better understanding of the problem and a deeper understanding of the tasks involved by focusing on experts on both ergonomic health and healthcare work.

3.1.2 Participants

During January and February 2016, four interviews were carried out. Two (Dutch) SME's in the field of healthcare, proper execution of tasks, and safe work environments for nurses were each interviewed twice.

3.1.3 Materials

Two structured interviews were conducted with each expert, the first focused on formalising the thought processes involved throughout the execution of the tasks, and the second focused on identifying possible pitfalls and other reasons for why nurses, albeit unconsciously, endanger their own health (see Figure 1). In the first interview we provided example problems which had to be resolved properly and the interviewee had to explain his/her reasoning behind each decision. For the second interview, two different techniques were used:

- a card-sort task of the different categories of tasks based on physical strain, endangering the patient, and decisions being context dependent (separately); and
- a free-recall on possible pitfalls and other reasons for why (and how) nurses endanger their own health, per task category. The recall was semi-structured by providing three categories to classify their reasons with; knowledge related, behaviour related, and logistics related.

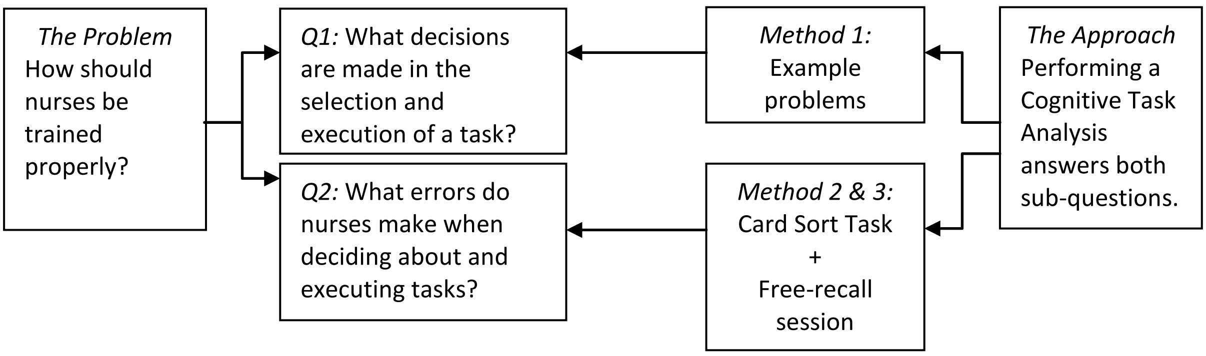

Figure 1: The problem and its sub-questions and their corresponding approaches for finding the answers.

3.2 Results

Providing a detailed overview of all the results of the CTA is beyond the scope of this paper, so instead we present only those findings that are relevant for the goal of the paper:

- Context is everything

An interesting finding about the tasks themselves was the amount of caveats and different constraints that could influence how a task has to be executed. Although the general description of each task (as shown in instructional videos provided by one of the experts) gives a clear overview of the steps of a task, it lacks the complexity of an everyday working environment. For example, some tasks require a room in which the patient's bed can be moved around freely in order for the nurse to be able to maintain a proper posture. However, both experts stated that this often isn't the case in the real world. In such cases, if the nurse wants to maintain a healthy working posture, a broader set of possible correct solutions has to be known. Other examples such as the emotional and mental state of the patient or tubes and other wires hanging around the bed can increase the complexity of a task considerably.

- Misbehaviour mostly due to incorrect routine
  The card sort task and free-recall session revealed that the incorrect working postures were not necessarily a consequence of lack of knowledge on the nurses' part. Instead, it is the working routine created through years of experience that causes most nurses to develop their chronic injuries. When the symptoms finally surface it is nigh impossible for the nurse to trace what aspect of her working routine has caused the injury.

3.3 Summary

The results of the CTA showed that our initial problem definition of teaching nurses proper working postures would not solve the actual problem, as the problem is more related to behaviour than knowledge. The required behavioural change was not something a serious game alone could do, as it would require manual, real-life training as well. Instead, we focus on the first finding, which shows how complex a task can become given its contextual aspects.

The diversity of problems that nurses have to solve highlights a need for knowing when to use what aiding tool or manual skill to perform the task, where the correct choice depends on the context. To do so effectively, nurses would need a mental 'toolbox' suited to perform the same task in a multitude of ways. Thus the CTA not only provided a more thorough understanding of the tasks a nurse has to do, but it also deepened our understanding of the problem. To solve this problem, we structured the knowledge and made an instructional design using the four component instructional design method (4C-ID).

4. Four Component Instructional Design

In an effort to create an instructional design method to teach complex cognitive skills, van Merriënboer, Jelsma, and Paas (1992) identified four essential interrelated components:

- Learning Tasks: concrete, authentic, whole-task experiences organized in specific task classes;
- Supportive information: information that is supportive to the learning and performance of non-recurrent aspects of learning tasks;
- Just-In-Time Information: information that is useful/required for the learning and performance of recurrent aspects of learning tasks; and
- Part-task Practice: smaller practice items for recurrent aspects of a larger complex skill.

Since its introduction, the 4C-ID model has proven itself as a valid approach for creating effective interactive training tools that can be used to teach complex skills (van Merriënboer, Clark and De Croock, 2002; van Merriënboer and Kirschner, 2012). Sadly, it is beyond the scope of this paper to go into detail as to how this method is normally applied.

Because it has been extensively used and validated, it can be considered an interesting option for educational game design, as the frameworks that have been developed for this field lack the empirical evidence and maturity that the 4C-ID model has. The documentation available for this model also makes its application less subjective, providing a formal, more universal, approach to designing the game. Using the ten steps provided by Van Merriënboer and Kirschner (2012), one technically has a mold for the game's level design as each of the four components and their steps work towards creating a piece of a level (e.g. the information types represent instruction and feedback).

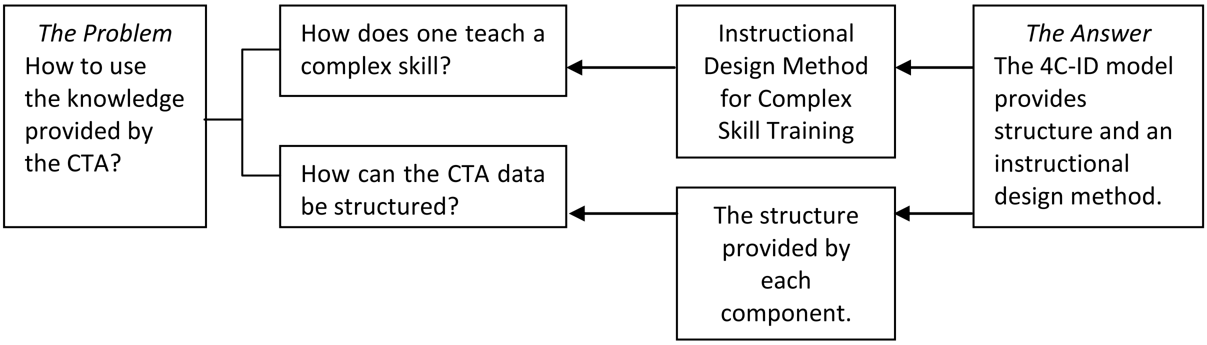

Figure 2: An overview of how the problem and its subsequent question (on the left) are answered by the 4C-ID model (on the right).

This model's validity, accompanied by previous educational game design research showing its usefulness in conjunction with CTA (Van Rosmalen et al., 2013), gives a strong incentive to use this model as the instructional design framework for our game. Not only does it fit naturally with CTA by providing its user with an approach to teach the complex skill one has researched, it also provides the instructional structure for the knowledge in the form of the four components (see Figure 2). Before we discuss the influence of the 4C-ID model on the game's design, the game itself will be discussed briefly.

4.1 The Game – Can You Help?

The game used in this case study closely resembles another serious game called *SURGE: Fuzzy Chronicles*. This game, designed by Adams and Clark's (2014) team, teaches the user about Newton's laws of physics. It does so by asking the user to plan out the actions of a space ship, with the goal of having the space ship move from point A to point B. The user has to select actions and their appropriate amount and direction of force and press a button to launch the ship, which will execute the actions in the order the user has programmed them. The game play requires the user to predict what the chosen actions will do and choose the right strategy accordingly, demanding a certain amount of thought to be put into the user's actions.

This 'planning' mechanic, in which the player creates a timeline of actions that the avatar will execute, is also used in the game 'Can you help?'. Instead of a space ship, the user has to make a human avatar move an object from one point to another. The actions range from manual actions such as lifting, carrying, and pushing, to actions which involve aiding tools such as a wheelchair or an electricity-powered hoist.

Before the application of the 4C-ID model, there was only the idea of this game play and a simple prototype showcasing it. What still lacked were proper level design, feedback, and an idea for progression. All three of these aspects are fleshed out using the guidance and ideas found in the 4C-ID model, as described in the next section.

4.2 Application to Game's Design

Using the knowledge gained from the CTA, the application of the 4C-ID model leads to the following global design decisions:

- Whole-task, authentic experiences
  The game's challenges will incorporate examples of problems that can occur in a nurse's everyday work. As the CTA has shown, the less ideal situations not explained in the regular documentation provide a challenge that the nurses not necessarily now how to deal with correctly (from an ergonomic point of view). The 4C-ID method helped to guide the sequencing of these experiences and the selection of when to present what situation and why.

- Maintaining difficulty through variety
  Teaching nurses a 'toolbox' approach to solve unexpected ergonomically problematic situations requires a variety of situations in which the learnt tools are useful and effective. Key here is to use the structure provided by the 4C-ID model to introduce new tools through the task class structure and

managing the difficulty by providing a variety of situations within a task class. The task classes help to train the recognition of situations in which the tool should be used, as initially one task class will have solutions requiring one tool. The variety within each task class supports the recognition training more thoroughly by providing a good fit in difficulty and ensuring that the tool is used in multiple scenarios.

- Supportive and procedural information focused on strategy
  As the 'toolbox' approach requires a thorough understanding of the problem situation, the focus of supportive and procedural information should be on finding the knowledge required for tool selection. Supportive information will deal more closely with the tools and when to use them, whereas the procedural information will help the user think about how to use the toolbox itself. For the latter, an example would be to not necessarily focus on a specific error made whilst playing, but to provide feedback on the more general approach taken by the user.

Part-task training was found to be less necessary to include, as this would come down to learning more general facts about ergonomics. This knowledge, as also found in the CTA, is often already known by caretakers. The task this game wants to provide repeatedly, the recognition of problem situation's contextual factors, is already being taught by the task classes themselves.

4.3 Summary

Applying the 4C-ID model to the game design process of *Can You Help?* helped to structure the underlying instructional design of the game. Not only did it provide a suitable structure to apply the CTA data to, it also provided an instructional lens for some of the core game elements (e.g. level design, progression, feedback). Figure 3 provides an overview of how the results of the CTA and its subsequent structuring using the 4C-ID model have led to the design of these core game elements.

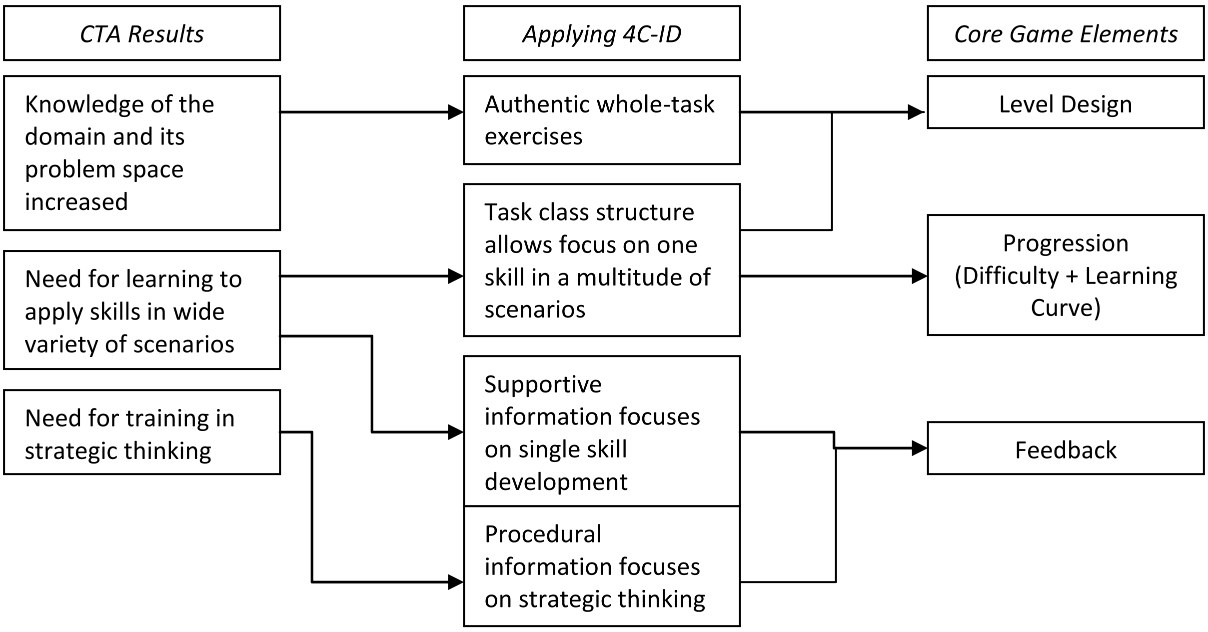

Figure 3: Overview of the influence of each applied formal method on the design of three core game elements.

5. Discussion

This section will discuss the evaluation of the two methods described above and will provide some advice on their usage.

5.1 Cognitive Task Analysis – Understanding the 'what'

CTA provides a vast set of tools for one goal: finding out *what* your game requires for it to achieve a certain goal. As the CTA consists of a large selection of different methods, choosing the right ones can be quite the challenge. The following experiences should help other researchers with this problem:

- Know what you need
  Not knowing the slightest about your problem will make it difficult to select the right CTA methods. Therefore it can be useful to have some exploratory interviews with experts to get a feel for what methods to choose. For example, we found out that our problem was mostly a behavioural problem, as nurses had developed their own, often wrong, working routine. A CTA will not provide information on how to persuade nurses to change their behaviour, as it is a 'cognitive' task analysis; implying that it focuses mainly on the underlying thought processes and decisions being made during the task. While we did gain useful knowledge on the thought processes and decision making of the tasks, the problem required different information.

- Level of detail
  How low-level does the extracted information have to be? Different tools provide different levels of detail. For example, during our CTA we also had an observation session in which the behaviour of one of the experts was observed during her shift. The reason to do so was to see whether things had been 'forgotten' during the interviews. The level of detail from this observation turned out to be too high. The things that the expert did not mention, such as the exact stance for a certain task, were too concrete and physical to add into the eventual design.

- Scope of the game
  Another important consideration has to do with establishing the intended scope of the game. By analysing the problems in more detail through a CTA, we discovered that the problem space is actually quite large and consists of many different aspects. We could focus purely on the specific behavioural aspects of a small selection of tasks and create a reasonably sized game around it. After determining the problem space through the CTA, one really needs to choose the most salient goals to include into the (final) game.

Another interesting aspect of using CTA for preliminary research is the influence it has on potential design space. Design space describes the breadth of options and tools a designer has to create a game; a basic example would be the general game play style (e.g. a first-person shooter or top down strategy game). In educational game design this design space is often quite small due to what a game aims to achieve in terms of learning. We reason that CTA diminishes this space even further, as it provides a more detailed, and thus narrowed down, understanding of what the user needs and thus what the game needs. However, we also believe that having the data of the CTA at ones disposal increases the *use* one can get out of this limited design space (i.e. it is easier for the designer to understand the consequences of the design choices with relation to the overall goal). A better understanding of the domain allowed us to be more creative within the constraints of the problem.

5.2 Four Component Instructional Design – Providing the 'how'

After understanding the details of a complex skill, the next step would be to design the instructional aspects surrounding that skill. 4C-ID provides the structure to implement the CTA data in, creating an instructional design which can in turn be translated to game design elements. While designing the instructional foundations of the game, it is important to keep track of the following things:

- Task classes and their depth
  Early on in our use of the 4C-ID model, we used all the information the CTA provided to create a wealth of task classes. These task classes would cover all the possible exceptions and combinations of contextual cues. This quickly turned out to be a bad approach to using the 4C-ID model, as an enormous amount of task classes would have had to be made to cover everything. Instead, we changed our approach by looking at the commonalities between tasks and creating task classes which dealt with teaching about these commonalities, reinforcing recognition in the user.
  The guideline here is to make clear learning goals per task class and to make sure that each task class' main focus is to achieve that learning goal.

- Translating 4C-ID components to game design
  Although the 4C-ID model provides an instructional structure for the CTA data, we experienced some difficulty with translating all its components to game design elements. As also seen in CHERMUG's design, the part-task training component was left out. It is important not to force the structure onto your game's instructional design, but to identify the aspects of the model your game benefits from the most.

5.3 The Translation Conundrum - Answering the 'Why'

There is one troublesome part of using the 4C-ID model for a game's design, which is the fact that the step from instructional design to game design is not covered in the 4C-ID literature. This would mean that all the effort put into having a formalised approach to game design could be lost during this translation. Huang and Johnson (2009) provide some guidance in linking these two design processes, as their work describes the relationship each game element has with each of the four components. However, their work is not based on empirical data, and only provides insight into the *instructional* aspect of game design.

We believe that Carvalho et al.'s (2015) activity theory-based framework (ATSMG) could provide an interesting follow-up to the 4C-ID model, allowing one to provide reasoning behind the translation from instructional design to game design. Although the framework is still young, its design is meant to provide insight in answering *why* each game design decision was made. Future research could look into ATSMG, and whether it can use the instructional design resulting from the 4C-ID model as justification for game design choices.

6. Conclusion

This paper has presented an overview of the use of two formal methods: one for analysing complex skills and one for teaching complex skills. Both of these have been used for a case study on training nurses in general problem solving behaviour using an educational game. The preliminary research to identify the problem domain, pitfalls, and decision making processes in experts was done using CTA. With two rounds of interviews and three methods, our initial problem definition was discarded after identifying the real underlying problem. The design process that followed was structured using the 4C-ID model, aiding the instructional foundations of the game's design as well as providing a skeleton for the CTA data.

We believe that the way these two processes are handled now: the preliminary research with its resulting problem definition & learning goal and the design process with its resulting core mechanics & level design, could be done more universally by using formal methods that have stood the test of time. Although these methods were not created with educational game design in mind, we hope to have shown with this paper that the possibilities are worth exploring.

7. References

Adams, D. M., and Clark, D. B. (2014) "Integrating self-explanation functionality into a complex game environment: Keeping gaming in motion", *Computers and Education*, Vol. 73, pp 149–159.
Bellotti, F., Kapralos, B., Lee, K., Moreno-Ger, P. and Berta, R. (2013) "Assessment in and of serious games: an overview", *Advances in Human-Computer Interaction*, Vol. 2013, No. 1.
Boyle, E., Van Rosmalen, P., MacArthur, E., Connolly, T., Hainey, T., Johnston, B., Moreno Ger, P., Manjón, B., Kärki, A., Pennanen, T., Manea, M., and Starr, K. (2012) "Cognitive task analysis (CTA) in the continuing/higher education methods using games (CHERMUG) Project", In *6th European Conference on Games Based Learning,* p.p. 63, Academic Conferences Limited.
Bronkhorst, B., Ten Arve, A., Spoek, M. and Wieman, D. (2014) "Gezond werken in de zorg. Onderzoek naar fysieke en psychosociale arbeidsbelasting onder zorgmedewerkers", Erasmus University Rotterdam, Stichting IZZ, Apeldoorn.
van Bussel, R., Lukosch, H. and Meijer, S. A. (2014) "Effects of a Game-Facilitated Curriculum on Technical Knowledge and Skill Development", In *Frontiers in Gaming Simulation* (pp. 93-101). Springer International Publishing.
Carvalho, M. B., Bellotti, F., Berta, R., De Gloria, A., Sedano, C. I., Hauge, J. B., Hu, J. and Rauterberg, M. (2015) "An activity theory-based model for serious games analysis and conceptual design", *Computers & Education*, Vol. 87, pp 166-181.

Clark, R. E., Feldon, D., van Merriënboer, J. J., Yates, K. and Early, S. (2008) “Cognitive task analysis” *Handbook of research on educational communications and technology*, Vol. 3, pp 577-593.
Clark, D. B., Tanner-Smith, E. E. and Killingsworth, S. S. (2015) “Digital games, design, and learning a systematic review and meta-analysis”, Review of educational research, doi: 0034654315582065.
Clark, D. B., Virk, S., Sengupta, P., Brady, C., Martinez-Garza, M., Krinks, K, Killingsworth, S., Kinnebrew, J., Biswas, G., Barnes, J., Minstrell, J., Nelson, B., Slack, K. and D’Angelo, C. (2016) “SURGE’s Evolution Deeper into Formal Representations: The Siren’s Call of Popular Game-Play Mechanics”, *International Journal of Designs for Learning*, Vol. 7, No. 1.
Degens, N., Bril, I. and Braad, E. (2015) “A Three-Dimensional Model for Educational Game Analysis & Design”, presented at *Foundations of Digital Games 2015*, Monterey Bay, CA.
Huang, W. D. and Johnson, T. (2009) “Instructional game design using cognitive load theory”, *Handbook of research on effective gaming in education*, Vol. 3, pp 1143-1165.
Mayer, I., Bekebrede, G., Harteveld, C., Warmelink, H., Zhou, Q., Ruijven, T., Lo, J., Kortmann, R. and Wenzler, I. (2014) “The research and evaluation of serious games: Toward a comprehensive methodology”, *British Journal of Educational Technology*, Vol. 45, No. 3, pp 502-527.
van Merriënboer, J.J.G., Jelsma, O. and Paas, F.G.W.C. (1992) “Training for reflective expertise: A four-component instructional design model for training complex cognitive skills”, *Educational Technology Research and Development*, Vol. 40, No. 2, pp 23–43.
van Merriënboer, J. J., Clark, R. E. and De Croock, M. B. (2002) “Blueprints for complex learning: The 4C/ID-model”, *Educational Technology Research and Development*, Vol. 50, No. 2, pp 39-61.
Van Merriënboer, J. J. and Kirschner, P. A. (2012) *Ten steps to complex learning: A systematic approach to four-component instructional design*. Routledge.
van Rosmalen, P., Boyle, E. A., Nadolski, R., van der Baaren, J., Fernández-Manjón, B., MacArthur, E., Pennanen, T., Manea, M. and Star, K. (2013) “Acquiring 21st Century Skills: gaining insight into the design and applicability of a serious game with 4C-ID”, In *Games and Learning Alliance*, pp 327-334. Springer International Publishing.
van Rosmalen, P., Boyle, E. A., Van der Baaren, J., Kärki, A. I. and Del Blanco Aguado, A. (2014) “A case study on the design and development of mini-games for research methods and statistics”, *EAI Endorsed Transactions on Game Based Learning*, Vol. 14, No. 3.