A Scoping Review of Heuristics in Videos Games Research: Definitions, Development, Application, and Operationalisation

ANNE OZDOWSKA, PENNY SWEETSER, and MAHSUUM DAIIANI, The Australian National University, Australia

Heuristics present a cheap and effective way of evaluating usability. However, in video games, evaluating unique player experiences that are dependent on individual preferences and abilities presents a challenge that goes beyond usability. Video games are more than just functional software, so games heuristics have been adapted to help examine functionality and experience. This paper reports on how papers published in the ACM Digital Library between 2012 and 2022 develop and apply heuristics in video games research. We found that heuristics are often used outside their intended purpose of being used in an expert evaluation. Instead, they are used as survey instruments, interview guides, codes for thematic analysis, and as design guidelines. This research contributes to HCI and video games research by distinguishing the terms design guidelines and design principles from heuristics. We make recommendations for researchers around developing heuristics and conducting video game heuristic evaluations. We propose a method for operationalising heuristics and make recommendations for the implementation of heuristics to improve the quality of video game heuristic reviews.

CCS Concepts: • **Human-centered computing** → **Heuristic evaluations**; • **Software and its engineering** → **Interactive games**.

Additional Key Words and Phrases: Heuristics, Video Games, Design Guidelines, Design Principles

ACM Reference Format:
Anne Ozdowska, Penny Sweetser, and Mahsuum Daiiani. 2023. A Scoping Review of Heuristics in Videos Games Research: Definitions, Development, Application, and Operationalisation. *Proc. ACM Hum.-Comput. Interact.* 7, CHI PLAY, Article 389 (November 2023), 23 pages. https://doi.org/10.1145/3611035

1 INTRODUCTION

Over 30 years ago, Molich and Nielsen [29] developed a set of nine usability heuristics to scaffold the process of evaluating user interfaces. Heuristic evaluations were traditionally used to examine the usability of software and were structured using a deficit model, in which only heuristic violations were recorded [32, 34]. However, video games embody a complex combination of usability, ergonomics, cognitive load, affect, and motivation [61, 63, 66]. As such, developing and using heuristics that examine video games presents a unique challenge that goes beyond usability alone. Heuristics for video games need to evaluate usability but also assess how effective the game is in providing a desirable experience.

Despite the long history of heuristics in human-computer interaction, there is a lack of dedicated research on heuristics in video games. In the field of video games, heuristics have been developed to assess a range of aspects including enjoyment [63], usability [42], gameful design [66, 67], playability

Authors' address: Anne Ozdowska, anne.ozdowska@anu.edu.au; Penny Sweetser, penny.kyburz@anu.edu.au; Mahsuum Daiiani, mahsuum.daiiani@anu.edu.au, The Australian National University, Canberra, Australia.

2573-0142/2023/11-ART389
https://doi.org/10.1145/3611035

[4, 8, 9, 27, 45], learning and pedagogy [14, 17], augmented reality [1, 16], game approachability [11], social aspects [37], and flow-like experiences [22]. However, the way these heuristic sets are applied in academic games literature tends to deviate from the original heuristic evaluation methodology set out by Nielsen and Molich [29, 34], where results from heuristic evaluations conducted by between three to five experts are aggregated and the severity, criticality, and frequency of issues is documented [46, 47].

Although there is a substantial body of literature on heuristics and video games, there is a lack of meta-research on the practice and best practice of designing and applying heuristics to games. Further, as there is limited published work that describes how, or even whether, heuristics are used in the games industry. This research focuses on academic publications. In this paper, we investigate the following research questions: **RQ1: How are heuristics designed and used in video games research? RQ2: How does this compare to best practice methods? RQ3: What lessons can we learn?** To address these research questions, we conducted a scoping review to investigate the design and application of heuristics in video games research over the last decade (2012-2022). We found that games researchers often use heuristics outside their traditionally intended purpose as part of an expert evaluation. Rather, games researchers tend to adapt heuristics and the way in which they are used to better meet their needs. Heuristics were often used to confirm the presence of an experience or functional aspect of a game, as opposed to finding violations. While heuristics are intended to be used as part of an expert evaluation, in video games research, they have been used as survey questions [1, 2, 19, 26, 41, 48, 51], focus group questions [5, 65], thematic analysis codes [18, 61], and as design guidelines [64]. Given this diverse use of heuristics to evaluate and design games, we also identify 19 video game heuristic sets published since 2004 and compare the way heuristics in these sets are constructed against a reference point best practice methodology (RQ2). We found that video game heuristics were rarely developed or applied in line with the reference point we selected as a best practice approach. The lack of consistency in employing a formal heuristic evaluation process in video game research could suggest that heuristics, or a formal evaluation approach, does not meet the needs of video game researchers.

Our review contributes to the field of HCI, video games, and heuristics research by identifying the gap between what we consider to be best practice methods for heuristic evaluations and the way in which heuristics are being used *in the wild* in games research. We believe that this comparison will lead to a better understanding of the practical uses and development strategies of heuristics in video game research. We challenge the view that heuristic evaluations should be limited to only identifying heuristic violations and suggest that video game heuristics can be be employed in a way that facilitates confirming the presence of experiences and functionality, as well as seeking to identify design problems and inform the design process. As a result of our findings from RQ1 and RQ2, we make five recommendations based on the lessons learned (RQ3) for reference when using heuristics for video game evaluations. These recommendations aim to disambiguate the way in which heuristics are reported in the literature. Further, we propose a method for operationalising heuristics to be multi-functional tools that can be used in the design, problem identification, and experience evaluation of video games.

2 DEFINING HEURISTICS IN GAMES RESEARCH

The terms “heuristics” and “heuristic evaluation” are described and used in a variety of ways in HCI literature. As there was some ambiguity around the way both terms were reported in the video game literature, we discuss each term in this section to provide clarity on our interpretations of their traditionally intended meanings and to provide a foundation for comparing the way in which heuristics are currently used in academic video games research.

2.1 Heuristics

Nielsen and Molich [34] refer to their original nine usability heuristics as being "principles which are generally recognized in the user interface community" (p.250). The term *heuristic* and what constitutes a heuristic was not very clearly set out in the HCI literature however. Heuristics are articulated in heuristic sets as semantic statements that reflect recognised principles. They should be constructed to strike a balance between specificity and generality and used to provide a basis for reviewers to evaluate usability [46, 48, 49]. Traditionally, heuristics have been designed to be used as part of an expert evaluation [13, 34, 42, 46, 49]. We consider that the traditional interpretation of heuristics to be *predefined semantic statements used in an expert evaluation.*

2.2 Heuristic Evaluation

Traditionally, the purpose of a usability *heuristic evaluation* is to *uncover problems with the design of user interfaces* [31, 32, 34]. Nielsen and Molich [34] state that a "heuristic evaluation is an informal method of usability analysis where a number of evaluators are presented with an interface design and asked to comment on it" (p.249). While this definition doesn't refer to identifying problems, the method they present to collect the comments is based on a deficit model and the collation of usability problems. In the field of video games, Tondello et al. [66] and Pinelle et al. [42] describe a similar problem focused approach.

The HCI literature appears to generally support the view that heuristic evaluations should be performed by three to five experts using a predefined set of heuristics [13, 34, 42, 46, 49]. While Nielsen [32] proposes that the severity of each problem is identified, Quiñones and Rusu [46] suggest that in addition, issues identified should be considered based on their frequency and criticality [46, 47]. Although, the severity, frequency, and criticality of problems are not always reported [53], there is a general consensus that the traditional objective of a usability heuristic evaluation is to identify faults. Problems identified from each evaluation should then be aggregated to achieve the highest breadth and depth of potential design issues [34]. We consider that the traditional intention of heuristic evaluations is to *uncover problems with a design.*

2.3 Heuristic Development

We found that the processes used for the development and construction of heuristics varied, and that this led to questions about the integrity of some heuristics sets [36, 42]. Pinelle et al. [42] state that most "heuristics are primarily based on literature reviews or author introspection, rather than on detailed information about design problems that commonly occur in games" (p.1454) and Paavilainen [36] raise further concerns about the lack of validation in the development of some heuristics sets. To address the integrity of the heuristic development process, Rusu et al. [49] proposed a six-step method. The six-step development process includes the following stages: exploratory, descriptive, correlational, explicative, validation, and refinement. Rusu et al. [49] also provided a seven-item list for constructing heuristics which proposes that each heuristic should be reported by describing the following attributes ID, Name, Definition, Explanation, Examples, Benefits and Problems. We consider this method of development as a best practice heuristic development approach.

2.4 Heuristic Use

We found that there was a lack of clarity in the HCI literature around how heuristics should be used. Heuristics sets are often referred to as being *fault finding* checklists used for identifying problems with a design. Carter and Potter [6] define heuristics as being a checklist of issues to consider. Nielsen and Molich [34] and Pinelle et al. [42] suggest that heuristics are like "principles" that can be used to "find usability problems" [42, p.1456]. Tondello et al. [66] argue that heuristics

could be either principles or guidelines, but they affirm that the intended function of heuristics is to "identify usability problems" (p.246). Endsley et al. [13] also refer to heuristics being used to identify violations, but their focus shifts slightly away from problem identification. They suggest that heuristics are like "guidelines" that can predict a prototype's ease of use and could also be used to confirm positive aspects of a design. Overall, while heuristics tend to be reported in the HCI literature as being associated with problem identification, there is some confusion around their function in a heuristic evaluation as to whether it is purely related to identifying fault or whether they can also be used to confirm the presence of functionality and experiences and also guide video game design.

Murad et al. [30] provide another perspective on how heuristics should be considered and describe a heuristic as being a "conceptual design guideline that applies to an entire interaction" (p.3). In addition to not necessarily seeking to fault find, this view raises the issue of whether heuristics should be applied to aspects of a design, the design as a whole, or both. Similarly, Quiñones et al. [47] suggest using a 5-point Likert scale to assess each evaluator's overall perception of the degree to which the interface complies or does not comply with the heuristic. This perspective is directed at both confirming compliance and identifying faults but it also contrasts with the traditional view that heuristics should be limited to identify problems [27, 31, 42]. These views refer to heuristics being used as tools to consider a whole-of-game perspectives as opposed to being limited to identifying single usability faults.

Video game heuristics that have been developed to highlight problems [27, 42] still rely on the reviewer's reflection of a subjective play experience which may also include very positive experiences and this may impact on a reviewers ability to identify faults. We suggest that the evaluation of video games is different to other types of usability testing where functionality assessment is driven by a means-to-an-end motivation. For functionality driven software like word processing software, problem focused heuristics that assess whether functionality is present or absent seems suitable because of the reduced need to consider experiences. In video games, the impact of functional changes can have more than purely functional impacts and this can upset the challenge/skill balance. We argue that even when video game heuristics are designed to identify problems, the nature of video games as interactive and goal-directed software introduces a layer of human perception that should be considered in addition to problems.

Outside HCI usability research, in the field of mathematics, heuristics are reported to be applied and used more broadly. Polya [44] described heuristics as having a more general application and wrote that a "heuristic, as an adjective, means 'serving to discover' " (p.113). In this definition, heuristics are discussed in terms of inductive reasoning, with no emphasis on whether the purpose of the heuristic is to find fault or confirmation or whether it applies to the whole interaction or part. Rather, the purpose of a heuristic when used in this sense is to provide "provisional and plausible" (p.284) reasoning. This is an important difference in interpretation because, traditionally, in the field of usability, heuristics have been designed to only identify problems. We suggest that video game heuristics should take on this approach of being designed in a way that "seeks to discover" to allow for diverse use.

2.5 Redefining Heuristics

We propose that video game heuristics might require a more balanced approach in the way they are constructed so that their construction matches their use. We suggest that they may need to be constructed to highlight both compliance and failings in a design, while considering these aspects from both functional and affective points of view and potentially also from both a specific and whole-of-game lens. Our view is that **video game heuristics should be redefined as being multi-purpose evaluative statements that can be used to identify problems, provide design**

guidance, and examine the game playing experience. They should also be able to be used at various stages during the design and development process to provide valuable feedback to the researcher/s or development team.

3 SCOPING REVIEW

To address our first research question (RQ1), investigating how heuristics are designed and used in video games research, we conducted a scoping review of papers related to heuristics in video games published in the ACM Digital Library (DL) between January 2012 and April 2022. To conduct this scoping review we followed the PRISM methodology [38]. There were three main phases involved in this review: identification of records, screening, and reviewing. The third author conducted the first two phases and the first author conducted the third phase. Figure 1 provides a graphical representation of the described paper selection process.

The first phase involved identifying records in the ACM DL that met our search criteria. To capture as many papers as possible that relate to using heuristics in video games, the search string included terms that we considered could be commonly used along with heuristics, such as design principles, design guidelines, and design metrics. Further, to capture as many papers as possible that use heuristics in video games, similar terms including game, gamification, and gameful design were added to the search string. We used the following search string: **[[All: "heuristics"] OR [All: "design principles"] OR [All: "design guidelines"] OR [All: "design metrics"]] AND [[All: "game"] OR [All: "video game"] OR [All: "gamification"] OR [All: "gameful design"]].**

The search was split into two date ranges to make returned papers more manageable to review. Each range of 5 years was sorted by relevance to obtain the most relevant papers from each 5-year block. The first date range found papers published in a five year period between January 2012 and December 2016 (n = 2359) and the second search returned papers published between January 2017 and April 2022 (n = 2991). The searches were sorted by relevance and the first 500 from each search were downloaded and saved to a Mendeley library. In total, 1000 papers were reviewed.

The second phase of the scoping review involved screening the downloaded papers. To screen each paper, the third author reviewed each paper's title, keywords and abstract for relevance to video games and heuristic use. If the keywords were present, the third author then scanned the body of the paper for relevance by conducting a keyword search within the paper. Papers that mentioned the keywords in the introduction, discussion, conclusion, or contribution section were included if they reported on the development, improvement, or use of heuristics, design guidelines, principles, or metrics. There were 414 papers included on this basis.

The third phase involved reviewing the remaining 414 papers more closely. The first author read each of these papers in full and wrote notes in Mendeley on how each paper referenced the keywords. Further, each paper was tagged based on the way heuristics, design principles, or design guidelines were referenced, type of study, method used, platform, and participant group. Papers were included if they related to the application or development of heuristics for video games. The first author developed a spreadsheet to track and map details of the selected papers. Each paper was reviewed again before being included in the spreadsheet and details recorded about the study, method used, participants, platform, whether/which existing heuristics were used and how/if the heuristics were used. There were 29 papers that met the inclusion criteria for this scoping review on heuristics in video games. Papers that were excluded at this stage of the review either did not use the term *heuristic* and only referred to the other terms (design guidelines/design principles), used the term *heuristic* with reference only to machine learning or artificial intelligence, were not related to video games, or were not written in English. A list of selected papers is presented in Table 1. Papers that only referred to design guidelines or design principles were also recorded in another tab of the spreadsheet. There were no papers that referred to the term design metric. The

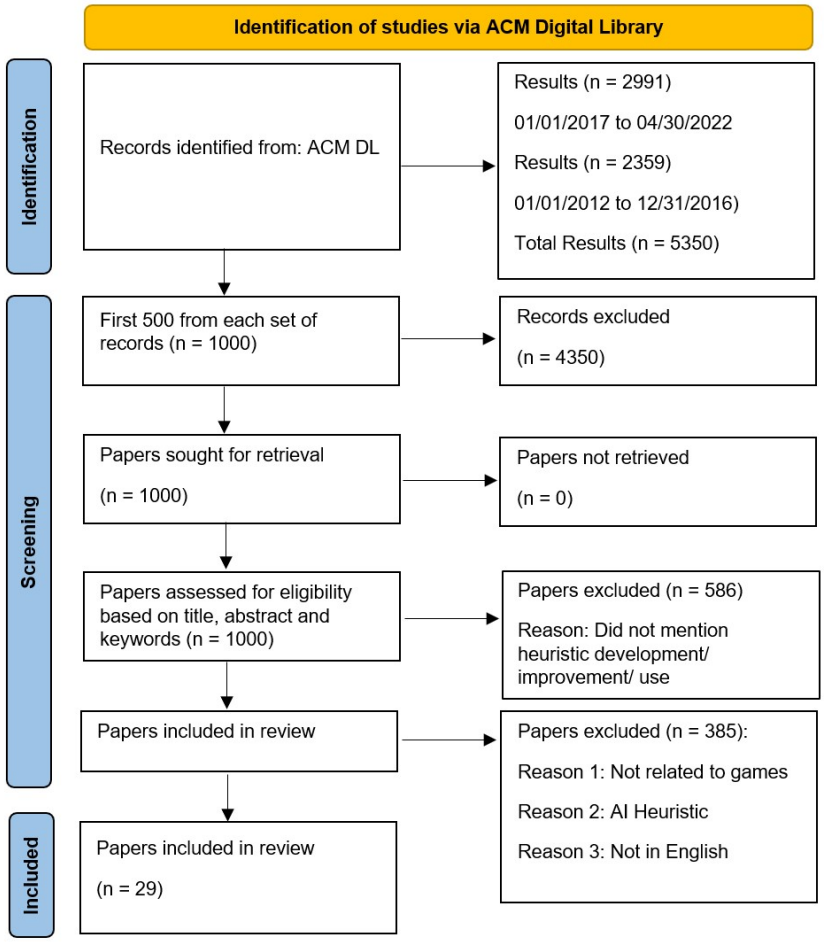

Fig. 1. Scoping review process for paper identification and inclusion [38]

same method was used to review and store details about each design guideline/design principle paper's type of study, method used, participants, platform, whether existing guidelines/principles were used and how/if the guidelines/principles were used. There were 143 papers that applied or developed design guidelines/principles. We included these terms in our original search to capture the widest possible range of papers that may have used heuristics. Papers in this set did not overlap with the heuristics papers. We have included a brief discussion about the 143 papers that were returned in our search that related to video games and either design guidelines or design principles in section 3.2.

For each of the 29 included heuristics papers that we reviewed, we identified which heuristic sets they referenced and found those papers. We then used a snowball method to identify any additional heuristics papers that reported on the development of new heuristic sets related to video games. We refer to these papers as being a Reference Heuristics Set. The date range of these papers was wider than the original 10 years that we had searched for in the ACM DL and spanned between 2004 and

2021. We evaluate how these papers structure their heuristics to identify a basis for comparing the ACM DL search results to and discuss this in section 5.2.

3.1 ACM DL Heuristic Paper Classification

For the scoping study, papers were classified based on the platform and domain that they were developed for and whether they developed a new heuristics or whether they used an existing heuristic. Table 1 sets out the platform, domain and the application of heuristics for each paper. There were six categories of video game platforms identified in the papers that met the selection criteria. These included Mobile or tablet games, Personal Computer (PC) games, Virtual Reality (VR) games, Wearable technologies, Console games, Unknown, and All. In one paper, it was not clear what the platform was used and a category of Unknown is reported. Where the paper was clearly applicable more generally, the category All was used. Papers were also categorised based on the main domain type that applied to the game. There were four domains. These included Education games, Commercial entertainment games, games designed for research purposes referred to by the category Research, and All. The All category was used where there paper applied to games more generally.

We also classified these papers based on how they applied heuristics. There were 10 papers that developed new heuristics [1, 6, 16, 19, 22, 23, 28, 37, 40, 67], but one of these papers reported on a set of *categories* and suggested that these categories would form the basis of new heuristics [40]. Table 2 sets out a more detailed analysis of the method used to develop each heuristic. There were three high level methods for developing new heuristics identified. These were qualitative, quantitative, or literature-only. The method used to inform the qualitative or quantitative approaches is also reported along with the participant group engaged in the development process (if any) and whether the heuristic was developed to confirm features in the game or find violations.

There were 19 papers that applied existing heuristics in a study [2, 5, 15, 18, 26, 41, 45, 48, 50–52, 56, 57, 60–62, 64, 66, 68]. One paper fell into both categories, as a new set of heuristics was developed and then used in a study [28]. Papers that applied heuristics in their method were classified based on the way the heuristics were used. Table 3 sets out how each paper reported on using existing heuristics. There were three categories identified to describe the approach used to apply heuristics. These were qualitative, quantitative and both. The method used for each approach is also reported. There were seven different methods for using heuristics identified. These included focus group, expert review, interview, design guidelines, design analysis, survey and thematic analysis. Participant group/s for each method are also reported along with how the study was framed for the participants. Framing of the study is reported in terms of whether the method sought to confirm features in the game or identify heuristic violations and whether the participants were asked to apply heuristics by reflecting on the whole game or to apply heuristics while reflecting on individual strengths or weaknesses.

3.2 Design Advice Paper Classification

We included the terms design principles and design guidelines in our search terms to capture as many papers that referred to heuristics as possible. Therefore, a large portion (N=143) of the papers read in full (N=414) did not report on heuristics, but instead reported on using or developing design guidelines, design principles and other terms including design frameworks, design recommendations, design considerations and design implications. We refer to these papers using the term *design advice* to separate them from heuristics papers, but also for ease of reference to collectively identify the group of design related terms that were not heuristics.

Out of the 143 papers that used one or more of the design advice terms, there were 105 papers that developed new design advice, only 11 used existing design advice, 18 both developed and used

Table 1. Papers selected and related platform and domains

Author	Year	Platform	Domain	Application
Sim [52]	2012	PC	All	Used existing
Tadayon et al. [64]	2012	VR	Entertainment	Used existing
Barcelos et al. [2]	2013	PC	Education	Used existing
Khanana and Law [26]	2013	PC	Education	Used existing
Paavilainen [37]	2013	All	All	Developed
Lucero et al. [28]	2013	Mobile	Commercial	Developed
Ponnada and Kannan [45]	2013	Mobile	Commercial	Used existing
Rodio and Bastien [48]	2013	PC	Entertainment	Used existing
Wodike et al. [68]	2014	Mobile	Entertainment	Used existing
Gale et al. [16]	2015	Wearable	Education	Developed
Carter and Potter [6]	2016	VR	Entertainment	Developed
Guo and Goh [18]	2016	PC	Education	Developed
Guo and Goh [19]	2016	PC	Education	Used existing
Serge et al. [50]	2016	PC	Entertainment	Used existing
Tondello et al. [67]	2016	All	All	Developed
Bunt et al. [5]	2017	Mobile	Education	Used existing
Sweetser et al. [60]	2017	PC	Entertainment	Used existing
Aultman et al. [1]	2018	Mobile	Research	Developed
Shu-Hui et al. [51]	2018	Console	Entertainment	Used existing
Gabriel Elías Chanchí et al. [15]	2019	PC	Education	Used existing
Jalife and Holmgrd [23]	2019	All	All	Developed
Pearson and Shaban [41]	2019	Unknown	Education	Used existing
Sweetser et al. [62]	2019	PC/VR	Entertainment	Used existing
Sweetser and Johnson [56]	2019	PC	Entertainment	Used existing
Tondello et al. [66]	2019	All	All	Used existing
Sweetser et al. [57]	2020	PC	Entertainment	Used existing
Sweetser and Rogalewicz [61]	2020	PC/VR	Entertainment	Used existing
Jalife et al. [22]	2021	All	All	Developed
Partlan et al. [40]	2021	All	All	Developed

the design advice, and 1 that updated design advice. There were 8 papers that made reference to developing or using design advice but it was not clear what the design advice was or how it played a role in the research. Additionally, of the 143 design advice papers, there were 36 papers that interchangeably used a combination of terms to refer to their advice.

4 SCOPING REVIEW FINDINGS

We present the findings from our scoping review in three parts. First, we report the 10 papers that developed new heuristics, second, we report on the 19 papers that applied a set of heuristics and third, we discuss the papers that referred to some form of design advice.

4.1 ACM DL Heuristic Development Papers

For the papers identified in the ACM DL scoping review development subset, there was no clearly dominant method of developing new heuristics. Heuristics were reported to be developed (at

Table 2. Heuristic development papers

Paper	Approach	Method	Participants	Considers
[1]	Qualitative	Interview/survey	8 Adults	Violate
[6]	Qualitative	Think aloud/survey	15 Adults	Violate
[16]	Qualitative	Expert review	Authors	Violate
[40]	Qualitative	Interviews	2 Experts	Confirm
[19]	Quantitative	Survey	39 University students	Violate/Confirm
[23]	Literature informed	-	-	Literature
[22]	Literature informed	-	-	Literature
[37]	Literature informed	-	-	Literature
[28]	Literature informed	-	-	Literature
[67]	Literature informed	-	-	Literature

least initially) based on the literature and researcher experience [16, 22, 37, 67] or by using the literature with some kind of experimental design [1, 28]. Both qualitative and quantitiatve methods were reported to be used when developing new heuristics. Of the 10 papers that developed new heuristics, four used a qualitative method to identify the new heuristics [1, 6, 16, 40], one paper used a quantitative method [19], and five papers relied on the literature [22, 23, 28, 37, 67]. Table 2 shows how each paper that developed new heuristics was categorised. There was a range of methods employed to develop the new heuristics, which included interviews, surveys, expert review, and think aloud sessions. Only one paper [16] referenced the heuristic development method set out by Rusu et al. [49]. Of the five papers that developed heuristics using participant input, three identified the new heuristics by asking participants to consider violations [1, 6, 16], one paper identified the new heuristics through seeking to confirm experiences [40], and one paper looked at both perspectives [19].

Most of the heuristic design papers identified in the ACM DL scoping review development set used a single method for the development of their heuristics. However, it has been recommended that mixed methods be employed to develop heuristics [47, 49, 53]. Sim et al. [53] argue that the use of a single method to develop heuristics could lead to important aspects being overlooked or bias in the heuristics and suggest that a mixed methods approach should be used. To develop a set of new heuristics for Computer Assisted Assessment systems, Sim et al. [53] used three approaches, they analysed the literature, conducted a survey and ran heuristic evaluation studies. Quiñones et al. [47] provide a similar set of steps that extend the work done by Rusu et al. [49] and their methodology for the development of new heuristics also incorporates a range of methods and points of data collection. They first propose that a scoping review should be conducted, second, that existing and potentially new experimental data should be collected and reviewed, and third, that the six steps set out by Rusu et al. [49] be followed. While Sim et al. [53], Quiñones et al. [47] and Rusu et al. [49] set out varying strategies for the development of heuristics, the common theme between each of their methods is that to develop reliable and valid heuristics, an iterative and mixed methods approach should be used. However, it should be noted that the steps set out by each of these authors relate to developing heuristics based on a deficit model and focus on the identification of problems. This approach may need to be adjusted for use in video games because heuristics are used to both confirm the presence of functionality and experience and to highlight problems or faults.

In examining how heuristics were developed for the papers included in our scoping review development subset, we noticed that video game heuristics were developed to assess different aspects of games, and to support the game development process. Heuristics sets tended to be developed to be used as qualitative tools. Most of the papers (8 of 10) that reported on new heuristics offered guidance on the aspects of games that they were designed to evaluate. There were three papers that referred to the heuristics as being suitable to evaluate a game [16, 19, 67], three papers that suggested that their heuristics could be used for game design [22, 23, 40], and two papers that suggested that the heuristics could be used for both design and evaluation [28, 36]. Only one of the papers suggested that the heuristic should be used quantitatively [19]. This contrasts with the papers in the scoping review subset that *applied* heuristics as many of these papers (12 of 19) used heuristics in at least some way as quantitative instruments (see section 4.2). We suggest that when heuristics are developed, the authors should clearly articulate how the heuristics are intended to be used and applied in practice. The outcome of this part of our review partially answered our first research question (RQ1), and we found that heuristics tended not to be developed using a predefined methodology like Rusu et al. [49] and that most only used a single method to develop their new heuristics.

4.2 ACM DL Heuristic Application Papers

The 19 papers in our ACM DL scoping review that reported on the application of heuristics in games research, used a mix of qualitative and quantitative methods, see Table 3. The main qualitative methods involved using heuristics in focus groups, interviews, and expert reviews. The main quantitative method was to use heuristics in a Likert scale survey. For the 19 papers that used heuristics, there were 7 papers that implemented heuristics using a qualitative method [5, 15, 18, 50, 52, 64, 66], 10 used a quantitative approach [2, 26, 45, 48, 51, 56, 60, 62, 62, 68], and 2 used a combination [41, 57].

Table 3 shows how each paper that used heuristics was categorised in terms of its methodological approach, participants, whether the study was framed to consider confirmations and/or violations of the heuristics and whether the heuristics were applied to the whole game or used to identify individual features. There were a number of methodologies reported in the subset. In addition to expert reviews, heuristics were used in surveys, interviews, focus groups, thematic analyses, as design guidelines, or as a basis to analyse a design. There were diverse groups participating in the research including children, teenagers, university students, adults, amateur players, developers, and experts. Further, heuristics were also applied to text from professional reviews. Heuristics were used in research to both confirm the presence of the experience or functionality associated with the heuristic, and to identify violations. There were 10 papers that used heuristics to confirm that a games was successful in affording specific experiences, 4 papers that sought to identify heuristic violations, and 5 papers that considered both perspectives. Only 1 paper attempted to report on the severity, frequency, and criticality of heuristic violations [68]. In most papers, participants were asked to reflect on the whole game when considering the heuristics, rather than to use the heuristics to identify individual violations or confirmations within one aspect of the game. Heuristics were used more to consider the overall experience of the game, rather than focusing on usability testing to identify problems. The outcome of this analysis provided a partial answer to our first research question (RQ1) and indicated that heuristics were not used exclusively in expert evaluations, but rather, in practice, were used in a variety of ways in video games research and with diverse participants.

Table 3. Heuristic application papers

Paper	Approach	Method	Participants	Considers	Reviews
[5]	Qualitative	Focus groups	20 University students	Confirm/Violate	Whole game
[15]	Qualitative	Expert review	Unknown	Confirm/Violate	Whole game
[18]	Qualitative	Interview	10 University students	Confirm/Violate	Whole game
[28]	Qualitative	Expert review	4 Experts	Confirm	Whole game
[50]	Qualitative	Expert review	1 Expert	Violate	Individual aspects
[52]	Qualitative	Design guideline	-	Violate	Individual aspects
[64]	Qualitative	Design guideline/Analysis	-	Confirm	Whole game
[66]	Qualitative	Expert review	5 Experts	Violate	Individual aspects
[41]	Both	Survey	12 Experts	Confirm/Violate	Whole game
[57]	Both	Expert review/Interview	2 Experts/1 Developer	Confirm	Whole game
[2]	Quantitative	Survey	30 University students	Confirm	Whole game
[26]	Quantitative	Survey	100 Children	Confirm	Whole game
[45]	Quantitative	Expert review	2 Experts	Confirm/Violate	Whole game
[48]	Quantitative	Survey	120 Experts/Amateurs	Confirm	Whole game
[51]	Quantitative	Survey	167 University students	Confirm	Whole game
[60]	Quantitative	Expert review	1 Expert	Confirm	Whole game
[62]	Quantitative	Expert review	1 Expert	Confirm	Whole game
[61]	Quantitative	Thematic analysis	10 Professional reviews	Confirm	Whole game
[56]	Quantitative	Expert review/Survey	12 Experts, 351 Adults	Confirm	Whole game
[68]	Quantitative	Expert review	20 Teenage males	Violate	Individual aspects

4.3 ACM DL Design Advice Papers

The majority of papers returned in our scoping review related to some form of design advice, whether it was design guidelines, design principles, design frameworks etc. While we can't make many generalisations about this diverse set of papers, a difference that was noted between heuristics and all forms of design advice was that heuristics tend to be reused more often in comparison to design advice. One reason for this could be that heuristics can be used as both high-level design support and as discrete yes or no tools to identify gaps in functionality or problems with aspects of user experiences. Only around 8 percent (11/143) of the design advice papers used an existing set of design guidelines/principles/recommendations etc and around 13 percent (18/143) of the papers that used some form of design advice, developed that advice as part of the same research. Conversely, all of the heuristics papers that used heuristics reused an existing heuristic. Also, even in the smaller set of papers that developed new heuristics, 90 percent (9/10) of papers either built on existing heuristic sets or considered other heuristics in their evaluation of previous work. This shows that heuristics appear to attract a higher rate of reuse when compared directly with the rate of reuse of design guidelines. However, it should be noted that this comparison is not necessarily directly like for like. Forms of design advice have a more limited range of options for reuse whereas heuristics have been used in a variety of ways and with multi-purpose uses and this may contribute to the ease of reusing heuristics in the design process over other types of design advice. Further, not all design advice may be used in academic publications. It is possible that there is industry adoption of this advice that is not captured through the academic publication process.

5 ANALYSIS OF HEURISTICS IN GAMES RESEARCH

After analysing the ACM DL scoping review papers, we identified three additional aspects related to using and developing heuristics for video games which were important to consider to provide a broader perspective to our scoping review findings and to provide a deeper perspective for consideration of our research questions. First, we noticed that some of the heuristics papers used heuristics in non-traditional ways and appeared to be suggesting that heuristics could be used in line with the way design guidelines or design principles might be used as part of a design process. We therefore attempt to disambiguate these terms. Second, the non-traditional use of heuristics often led to variations being made to the semantic construction of heuristics, so we examine a broader sample of video game heuristics to examine their construction. Third, we found that there was a lack of consistency behind the way video game heuristics were phrased. Some heuristics appeared to refer to seeking to identify problems whereas others seemed to refer to confirming experiences, some relate to the whole game and some related to discrete aspects. An analysis of the intent behind heuristic phrasing is also presented.

5.1 Heuristics Versus Design Guidelines/Principles

While there were only 29 papers identified from the search that met the inclusion criteria for our ACM DL scoping study, there were some papers that used the terms heuristics, design guidelines, and design principles interchangeably [3, 7, 21, 41, 54] and some papers indicated that their heuristics could be used as part of the design process [22, 23, 40]. For heuristic papers, the inconsistency in the way in which these terms were reported could stem from the underlying intent behind how they are used in practice. To further examine RQ1 and provide additional clarity, we provide an overview of the terms design principles and design guidelines to clarify our understanding of the differences between the terms and their uses as compared with heuristics.

Design principles are described by Palalas and Wark [39] as being "broad, high-level, generalized, universal recommendations that are based upon empirically-tested theories that can be

applied across various contexts". In other words, they are used as high level design rules that have some leeway around their application. In contrast, **design guidelines** are described by Palalas and Wark [39] as being "specific, practical, and testable criteria for how to best achieve design principles within particular contexts". Design guidelines offer a practical way to apply the design principles. While heuristics were originally described by Nielsen [31] as being used exclusively for expert review as part of a usability analysis, they are described more recently in the context of video games, as being useful during the game design and development process [42, 59, 66]. Heuristics were reported as being used as a tool to aid video game design [52, 64] or designed to be used as design tools [22, 58, 67], tools to both frame the desired high-level game-wide experiences [63] and localised experiences or problems [42]. In practice, video game heuristics appear to have an overlapping function with design guidelines, design principles and other forms of design advice.

Pinelle et al. [42] highlights this grey area when suggesting that their Game Heuristics "can serve as a set of design principles that can be used during the formative stages of game design and development" (p.1455) and that they can also be used by expert evaluators as part of a heuristic evaluation. There were two papers in our scoping set that reported using heuristics as design guidelines [6, 64]. Tadayon et al. [64] clearly articulated the connection between heuristics and design guidelines when they asserted that "social playability heuristics, such as SoPlay, can be used to evaluate and guide development of social interactions in games" (p.85). Carter and Potter [6] provided an interesting perspective on the use of heuristics as a design tool and used Nielsen's usability heuristics to develop a game that violated them. They provided a detailed description of how they used the heuristics to guide the development of a truly terrible and frustrating game. However, in doing so, they demonstrated the importance and validity of these heuristics, not only as a tool for evaluation but also as a design tool. Further, they showed a tangible outcome of the impact that violations could have on user experience. The outcome of these varying definitions demonstrates a lack of clarity between the terms.

The ambiguity around how heuristics should be used was also highlighted by Tondello et al. [66, p.255] through the statement "heuristic evaluation or usability inspections allow experts to evaluate a design based on a set of principles or guidelines (i.e., heuristics)". While this statement provides a clear explanation for the term heuristic evaluation, it also suggests that heuristics are equivalent to principles or guidelines, a view shared by Nielsen and Molich [34]. Describing heuristics using terms that have meanings in their own right in the field of HCI has led to some interchangeable use of the terms and potential confusion about how heuristics are intended to be used. It could be that the definition provided by Tondello et al. [66] was intentional however, and the overlap in terminology is intended to allow for heuristics to be used in a multi-functional manner.

The interesting dilemma is where or whether heuristics fit within the hierarchy of design advice related terms. Heuristics are reported as being used as a tool to aid video game design [52, 64] or designed to be used as design tools [22, 58, 67], tools to both frame the desired high-level game-wide experiences [63] and localised experiences or problems [42]. Heuristics, in practice, are used as multi-function design and evaluation tools. This extended use of heuristics has led to the development of some heuristics that are specifically constructed for this multi-function use [42, 59, 66] but not all sets of heuristics are constructed with a multi-function use in mind. To examine how heuristics sets have been designed, we compare the way in which they are constructed with a point of reference that we have selected as a "best practice" heuristic development approach (RQ2).

5.2 Heuristic Construction in the Reference Heuristic Set

In addition to the papers we found in the original ACM DL scoping review literature search, we identified 19 different sets of video game heuristics using a snowball method, the Reference set.

We analysed the construction of heuristics in the Reference set and compared them with the the heuristic construction method set out by Rusu et al. [49]. We use the method set out by Rusu et al. [49] as a basis for best practice, but we acknowledge that there may be other best practice heuristic development methods.

None of the 19 video game heuristic papers reported heuristics using all 7 attributes (ID, Name, Definition, Explanation, Examples, Benefits and Problems) and no heuristic reported on the *benefits* or *problems* attributes. There was only one paper identified that provided *examples*. Table 4 shows how each set of heuristics was constructed against the format suggested by Rusu et al. [49]. We support the inclusion of examples, benefits, and problems when reporting on new heuristics because providing these additional insights could help to facilitate better reviewer perspective taking and more balanced evaluation outcomes. The outcome of this analysis partially answered our second research question (RQ2) and provided the additional insight that video game heuristics were not designed and developed in line with the best practice exemplar we selected. Most sets of heuristics did present a name, definition and explanation of each of their heuristics but it was not clear by looking at the heuristics what methodology was intended to be used. Given the range of uses for heuristics that we reported in our scoping study, we recommend that when new heuristics are defined that researchers consider indicating how they are validated to be applied in practice.

5.3 Heuristic Phrasing and Interpretation in the Reference Heuristic Set

In our scoping study, we found that heuristics were used to both identify problems and to confirm the presence of experiences. We therefore chose to examine how video game heuristics were phrased. When we examined heuristics in the Reference set, we found that the phrasing of video game heuristics appeared to be inconsistent with some heuristics seeking to identify problems and some seeking to confirm experiences. We provide some examples of heuristics from Table 4, HEP [8], PLAY [9], Heuristic Evaluation for Games [42], GameFlow [58, 63], Networked Game Heuristics (NGH) [43], Playability heuristics for mobile games [27]. In these examples, participants applying heuristics in reviews appear to need to toggle between considering a game from the perspective of seeking to confirm or identify a problem for the whole game to considering whether a feature was present or not in one part of the game. The following list of heuristics provides examples of discrete phrasing that is nearly Boolean in nature. These heuristic examples could make fault finding clear as they are directed at one aspect of game functionality.

- Heuristic Evaluation for Playability (HEP) [8] - "The player can easily turn the game off and on, and be able to save games in different states" (p.1511).
- Games Usability Heuristics (PLAY) [9] - "Player does not need to read the manual or documentation to play" (p. 562).
- GameFlow [58, 63] - "players should feel a sense of control over the game shell (starting, stopping, saving, etc.)" (p.5).

Some examples of heuristics from the same heuristic sets that are more open to interpretation are presented in the following list. These examples offer more of an opportunity to understand if whole-of-game experiences have been achieved as opposed to directly identifying violations.

- Heuristic Evaluation for Playability (HEP) [8] - "The game is fun for the player first, the designer second and the computer third. That is, if the non-expert player's experience isn't put first, excellent game mechanics and graphics programming triumphs are meaningless." (p.1511).
- Games Usability Heuristics (PLAY) [9] - "The game gives rewards that immerse the player more deeply in the game by increasing their capabilities, capacity or for example, expanding their ability to customize." (p. 561).

Table 4. Video game heuristic structure based on Rusu et al. [49]

Heuristic name	ID	Name	Definition	Explanation	Examples	Benefits	Problems
Presence in consumer VR [6]	-	-	Y	Y	-	-	-
EGameFlow [14]	Y	Y	Y	-	-	-	-
FUSE [22]	-	-	Y	Y	Y	-	-
GAP [10]	-	Y	-	-	-	-	-
Gameflow [58, 63]	-	Y	Y	Y	-	-	-
Gameful Design [67]	Y	Y	Y	Y	-	-	-
Game Heuristics [42]	-	Y	Y	Y	-	-	-
PLAY [9]	Y	Y	Y	-	-	-	-
HEP [8]	-	Y	Y	-	-	-	-
Heuristics for Social Games [36, 37]	Y	Y	Y	Y	-	-	-
Mobile Augmented UI Heuristics [1]	Y	-	Y	-	-	-	-
Networked Game Heuristics [43]	-	Y	Y	Y	-	-	-
Pervasive GameFlow [24]	-	Y	Y	Y	-	-	-
Playability-Mobile Games [27]	Y	Y	Y	Y	-	-	-
RTS GameFlow [12]	-	Y	Y	-	-	-	-
Video Game Design Heuristics [23]	-	-	Y	Y	-	-	-
Wearable Augmented Reality [16]	Y	Y	Y	-	-	-	-
PLEX [28]	-	Y	Y	-	-	-	-
IGEF [65]	-	Y	Y	Y	-	-	-

- GameFlow [58, 63] - "Players should feel a sense of control over their characters or units and their movements and interactions in the game world" (p.5).

While these examples are only a small sample of heuristics, they highlight the diverse range of potential interpretations for each heuristic, from *yes/no* to requires an *individual interpretation, opinion and judgement* on the degree to which the game (often as a whole) achieves the heuristic. Further, these examples also raise questions about how video game heuristic reviews deviate from traditional usability heuristic reviews in that they are not only prompting reviewers to look for discrete violations but are also seeking to confirm the existence of overall experiences. Korhonen and Koivisto [27] argue that heuristic game reviews should be conducted by experts as a result of the need for the reviewer to have the appropriate knowledge to apply the heuristics and to provide feedback that reflects having "really played the game" (p.12). Being able to learn the game in order to examine it while interpreting the heuristics, implies that heuristic video game reviews should be conducted by "double experts" [31]. Nielsen [31] describes double experts as reviewers with domain and usability knowledge.

Some of the research that we reviewed in our scoping study engaged double experts to conduct expert reviews, but they tended to be used to report which heuristics were present or met by the game as a whole [15, 28, 45, 60, 62], without identifying individual violations. There was only one paper that engaged double experts in expert reviews to identify heuristic violations [66]. The lack of research that engages double experts in heuristic evaluations that follow a traditional method indicates that this kind of approach might not be meeting the needs of games researchers. This could be because traditional heuristic evaluation methods only seek to identify heuristic violations and do not examine player experiences. It could also be that finding double experts is challenging, and having mixed intentions for games reviews (both looking for violations and seeking to examine experiences) could explain why some researchers choose to utilise heuristics in their research in non-traditional ways and why some research actively encourages engaging a diverse range of participants in heuristic evaluations [56, 57, 60].

Due to the nature of games needing to afford enjoyment, or at least to afford a desired experience, meaningful heuristic video game reviews require a specialist and multi-layered approach. Selecting the appropriate players to conduct the review is a critical component of the quality of the heuristic review. Rodio and Bastien [48] found that different groups of players placed higher or lower degrees of relevance or importance on each heuristic in a set. They suggest that heuristics could be weighted depending on how each reviewer is categorised. However, in practice, it could be difficult to accurately categorise or evaluate reviewers and match weightings accordingly. Further, this process could end up unintentionally skewing the results. While it is common to see the categorisation of reviewers reported, it is not clear whether weighting the heuristic results based on reviewer type would provide additional value to a heuristic review. It might also only apply to reviews that were using a quantitative data collection method. Rodio and Bastien [48, p.93] make the point that "it is impossible (and counterproductive) to optimize a game in all its aspects. Rather, priorities should be identified, according to a certain type of attended game play". It could be inferred that there is a role for experts who conduct reviews to identify these priority areas and for researchers to pay particular attention to these priority areas when reviewing the evaluation data collected from diverse groups of players. We suggest that to obtain the best breadth of data from a range of reviewers that heuristics should be operationalised for use by diverse groups.

6 OPERATIONALISING HEURISTICS FOR GAMES - LESSONS LEARNED

In considering the way in which heuristics were reported to be constructed, developed and applied in the field of video games research, we suggest that video game heuristics may need to be modified

to better represent their multi-purpose uses. To address our final research question (RQ3), what lessons can we learn, we synthesise the outcomes from our findings and propose five strategies for operationalising video game heuristics. At this stage these strategies are preliminary suggestions and we plan to explore these further in future research. Our preliminary operationalisation strategies include using a development strategy when designing new heuristics, clearly articulating data collection methods and intended uses for heuristics sets, standardising the layout of newly reported heuristics, including an importance ranking for each heuristic and including at least some qualitative data as part of the heuristic evaluation. By using and developing heuristics in this way, we believe that heuristics can be better operationalised for use as multi-function tools that can be used in the design, problem identification, and experience evaluation of video games.

Table 5. Recommended standardised layout for reporting new heuristics

Category	Attributes [49]	Practical use note
ID	Definition	Single interaction/whole game/either
Name	Theoretical grounding	Intended to confirm or find fault
	Examples	Proposed importance ranking method
	Expected benefits	Ability for reviewer to add qualitative feedback
	Problems/Misunderstanding	Application method (e.g. Expert review)

Table 6. Proposed additional data collection when using heuristics in an evaluation

Category	Additional data collection
ID	One interaction/whole game/either
Name	Confirm/Find fault
	Importance ranking
	Qualitative text field

6.1 OS1: Development Strategy

The first operationalisation strategy (OS1) that we propose for video games heuristics is to ***use a formal and iterative development strategy***, including a robust validation process. However, we note that current heuristic development processes are designed to identify problems. We suggest that while the process of identifying problems is important for the heuristic development, desirable features and experiences should also be considered.

6.2 OS2: Intended Method of Application

The second operationalisation strategy (OS2) that we propose for video games heuristics is that ***new heuristic sets articulate the method in which the heuristics should be used***. Having this method set out by the researchers who develop the heuristics can assist those who use heuristics in their work to implement them as they were intended. For example, whether or not the heuristic has been designed to be used as part of an expert evaluation and/or used as a survey.

6.3 OS3: Standardised Layout

The third operationalisation strategy (OS3) that we propose for video games heuristics is that ***a standardised heuristic layout should be adopted when reporting on newly developed***

heuristics. Building on Rusu et al. [49], we recommend that this layout should clearly articulate the following: category, ID, Name, details of the heuristic's definition, theoretical explanation or reference to applicable usability principles, examples of heuristic violation and compliance, expected benefits associated with compliance, and any anticipated problems related to misunderstanding the heuristic [49] (see Table 5). In addition to the attributes identified by Rusu et al. [49], and evidenced by the practical implementation and uses of heuristics that we have described, we suggest that it be made clear whether the heuristic was developed to apply to the whole game, to aspects of the game, or both. We also suggest that reviewers using the heuristic be required to identify this when they respond to a heuristic. Table 6 sets out the additional data we suggest is collected when a reviewer responds to each heuristic.

6.4 OS4: Importance Ranking Scale

The fourth operationalisation strategy (OS4) that we propose for video game heuristics is that ***each heuristic should be given an importance ranking scale***. By referring to the importance of a heuristic, no assumption is made that there is a problem which may be the case when severity or criticality is reported. When heuristics are used in research, each reviewer should assign an importance ranking to the heuristic assessment. In addition, reviewers should be able to indicate whether they are identifying a violation or confirming the successful implementation for each heuristic. The collection of this additional data would enable researchers to determine whether a fault or confirmation is being reported.

6.5 OS5: Inclusion of Qualitative Feedback

Finally, the fifth operationalisation strategy (OS5) that we propose for video games heuristics is that ***responses to heuristics should allow for some qualitative feedback***. By allowing evaluators to provide qualitative feedback, additional insight can be obtained about individual interpretations or specific problems encountered. Obtaining these additional reviewer insights could improve the prospects of interpreting reviewer scores or feedback. We envisage that the collection of this extra information would take the form of additional options associated with each heuristic.

7 DISCUSSION

In this research, we conducted a scoping review to examine the way in which heuristics have been applied in video games research over the last 10 years (2012-2022). We found that the reported use of heuristics in video games differs from traditional usability research, which focuses on identifying problems [32, 34] and recording their corresponding severity, criticality, and frequency [46, 47]. We did not identify any video game papers that aligned with all aspects of a traditional usability heuristic evaluation or development process. Rather, video game researchers tended to use heuristics in non-traditional ways with diverse participants. These findings are consistent with research conducted by Sweetser [55] who found that only 9% of the papers that reported using the GameFlow heuristics used them in an expert evaluation. The most common use for the GameFlow heuristics was as a survey (around 59%). This is despite the GameFlow heuristics not being a validated survey instrument. Heuristics are not usually constructed to be used outside heuristic evaluations and this could account for some researchers reporting the need to modify the semantics of heuristics before using them in their research [2, 5, 25, 28, 41, 51].

We also found that the focus of video game heuristics and heuristic evaluations was more heavily oriented towards examining overall experiences, as opposed to identifying individual faults. This difference in approach may have led to the utilisation of heuristics in non-traditional ways. For example, Lucero et al. [28] modified a set of heuristics into a deck of tangible cards. The cards were used as a tool for the researchers to stimulate collaborative discussions and to prompt experts

to reflect on playful experiences in a game. The use of heuristics to confirm the presence of an experience also contrasts with the deficit approach set out by Nielsen [33, 34]. The shift away from heuristics being used to identify individual issues and rather to require players to reflect on the whole game was present in the majority of video game papers that reported using heuristics. Therefore, we suggest that video game heuristics, and their accompanying attributes, should be phrased to enable reviewers to consider multiple perspectives. We propose that heuristic evaluation methods for video games should enable reviewers to articulate exact aspects of the game that meet or violate a heuristic. This includes whether the heuristic confirmation or violation relates to the whole game or part of the game and the ranking of the perceived importance of each heuristic confirmation or violation.

Additionally, there appeared to be some overlaps in the terminology used to report on video game heuristics and we note that video game heuristic definitions do not clearly separate heuristics from design guidelines or design principles. This could be because video game heuristics can be used for the design and evaluation of games and not only to find problems. We also explored the difference between heuristics and other forms of design advice. We note that the rate of reuse of heuristics appears to be higher, even though there are many more sets of design advice. The need to produce generalisable findings is strongly driven through the academic peer review process and this could be contributing to the volume of published video game papers that produce design advice. O'Shea and Freeman [35] made the point, when referring to design frameworks, that "the ability to find practical information about individual frameworks is limited without prior knowledge" (p.2). They also add that access to this information is also often restricted due to the pay per view nature of many academic publications. Hodent [20] used stronger language when referring to the volume of academic design advice, lamenting that game developers "are plagued by too many examples of methodologies and theories" (p.21). The combination of a high volume of highly specific design advice and the pay-wall restrictions on the advice could contribute to the low rate of reuse of these resources.

We propose that heuristics attain a much higher rate of reuse in academic research because they are functional and intentional. Design guidelines are often the incidental outcomes of papers and not the primary purpose of the conducted work. Heuristics offer both a set of design criteria that can be utilised during the design process and a functional evaluation tool. We propose that to address the gap between theory and the practical use of heuristics in research, that heuristics are reported in an operationalised way. We suggest that the use of heuristics may be of a higher value to video game researchers than other types of design advice because heuristics tend to be multi-function tools that can be used for problem identification, design, and evaluation. It is acknowledged however, that these results are preliminary and we hope to explore a much larger sample of literature and industry heuristics use in future work.

7.1 Limitations and Future Work

There are several limitations of our work that need consideration. First, we only searched the ACM DL. In the field of human-computer interaction, the ACM DL captures the primary and premiere outlets for publication. However, it is possible that we missed some relevant publications that were not present in the ACM DL. Despite sorting the results by relevance, we did only review the first 500 from each search. Future work could take a broader approach to searching for papers on heuristics and could include publications from a variety of databases and review every paper returned in the searches. Second, our suggestions and recommendations have not been tested in practice and are only based on an analysis of the literature. Our next steps will be to follow our recommendations to develop and validate heuristics for video game research. Future research should seek to compare heuristic evaluation outcomes using a variety of evaluation methods. Third, this paper focused

on heuristics in video games research, rather than in the game development industry. While it is difficult to assess the rate of adoption of academic research in industry, the games industry has been criticised in academic literature for not adopting research into practice. In the field of games user experience, Hodent [20] suggests that for most game developers, examining user experience is either ignored or non-existent. Hodent [20] proposed that the reason for this could be due to the lack of "tools to guide us in advancing UX maturity" [20, p.10]. It could be that industry are not embracing academic research outcomes like heuristics because they are not presented in an operational *tool*-like format. Valuable future work could explore the usage of game design research outcomes in industry.

8 CONCLUSIONS

The main goal of a heuristic evaluation is to develop a preliminary understanding of an interface before it is released. Traditionally, heuristic evaluations were conducted to highlight aspects of user interfaces that breach accepted design rules. We conducted a scoping review of video game heuristics papers published in the ACM DL between 2012 and 2022. In response to our first research question (RQ1) we made two main findings. First, in video games research, heuristics are used to both identify problems and to confirm the presence of experiences often by requiring reviewers to reflect on the whole game. Second, heuristics are not always used in expert evaluations and are commonly reported to be modified so that they can be used in non-traditional ways such as for surveys, interviews and focus groups. Researchers might modify heuristics because they are traditionally framed to highlight individual problems and not to examine player experiences. Video game evaluations require consideration of usability, ergonomics, cognitive load, affect, and motivation and video game heuristics need to enable reviewers to report on each of these aspects from individual and whole-of-game perspectives. In response to our second research question (RQ2), we found that video game heuristics did not tend to be developed or used in line with the best practice. Based on the findings from our first two research questions, we make five key recommendations to better operationalise heuristics to reflect their multi-function use. We identify key learnings (RQ3) to improve the way heuristics are designed and used in future research and game development.

REFERENCES

[1] Andrew Aultman, Spencer Dowie, and Nelly Ann Hamid. 2018. Design heuristics for mobile augmented reality game user interfaces. *Conference on Human Factors in Computing Systems - Proceedings* 2018-April (2018), 1–5. https://doi.org/10.1145/3170427.3188580

[2] Thiago Barcelos, Geiza Costa, Roberto Mũoz, René Noël, and Ismar Silveira. 2013. Informal HCI: What may students learn from playability issues during a game design workshop? *ACM International Conference Proceeding Series* (2013), 116–119. https://doi.org/10.1145/2535597.2535613

[3] Laura Bartoli, Franca Garzotto, Mirko Gelsomini, Luigi Oliveto, and Matteo Valoriani. 2014. Designing and evaluating touchless playful interaction for ASD children. *ACM International Conference Proceeding Series* 1 (2014), 17–26. https://doi.org/10.1145/2593968.2593976

[4] Clara Bonillo, Teresa Romão, and Eva Cerezo. 2019. Persuasive games in Interactive Spaces: The Hidden Treasure Game. *ACM International Conference Proceeding Series* (2019). https://doi.org/10.1145/3335595.3335626

[5] Lance Bunt, Verona Leendertz, and A. Seugnet Blignaut. 2017. A heuristic evaluation of the design and development of a statistics serious game. *ACM International Conference Proceeding Series* 1, Episode 1 (2017), 1–7. https://doi.org/10.1145/3136907.3136940

[6] Lewis Carter and Leigh Ellen Potter. 2016. Designing games for presence in consumer virtual reality. *SIGMIS-CPR 2016 - Proceedings of the 2016 ACM SIGMIS Conference on Computers and People Research* (2016), 141–148. https://doi.org/10.1145/2890602.2890626

[7] Konstantinos Chorianopoulos, Michail N. Giannakos, and Nikos Chrisochoides. 2014. Design principles for serious games in mathematics. *ACM International Conference Proceeding Series* 02-04-Octo (2014), 2–6. https://doi.org/10.1145/2645791.2645843

[8] Heather Desurvire, Martin Caplan, and Jozsef A. Toth. 2004. Using Heuristics to Evaluate the Playability of games. *Conference on Human Factors in Computing Systems - Proceedings* January 2004 (2004), 1509–1512. https://doi.org/10.1145/985921.986102

[9] Heather Desurvire and Charlotte Wiberg. 2009. Game usability heuristics (PLAY) for evaluating and designing better games: The next iteration. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 5621 LNCS, July 2009 (2009), 557–566. https://doi.org/10.1007/978-3-642-02774-1{_}60

[10] Heather Desurvire and Charlotte Wiberg. 2010. User Experience Design for Inexperienced Gamers: GAP – Game Approachability Principles. In *Evaluating User Experience in Games*, Regina Bernhaupt (Ed.). Springer, 131–148. https://doi.org/10.1007/978-1-84882-963-3

[11] Heather Desurvire and Charlotte Wiberg. 2015. User Experience Design for Inexperienced Gamers: GAP—Game Approachability Principles. (2015), 169–186. https://doi.org/10.1007/978-3-319-15985-0{_}8

[12] Sha Ding, Ningjiu Tang, Tao Lin, and Shiyuan Zhao. 2009. RTS-GameFlow: A new evaluation framework for RTS games. *Proceedings - 2009 International Conference on Computational Intelligence and Software Engineering, CiSE 2009* (2009), 0–3. https://doi.org/10.1109/CISE.2009.5363526

[13] Tristan C Endsley, Kelly A Sprehn, Ryan M Brill, Kimberly J Ryan, Emily C Vincent, and James M Martin Draper. 2017. Augmented Reality Design Heuristics: Designing For Dynamic Interactions. , 2100–2104 pages. https://doi.org/10.4324/9781003011293-1-3-4

[14] Fong Ling Fu, Rong Chang Su, and Sheng Chin Yu. 2009. EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Computers and Education* 52, 1 (2009), 101–112. https://doi.org/10.1016/j.compedu.2008.07.004

[15] G. Gabriel Elías Chanchí, Wilmar Yesid Campo Muñoz, and Luis Freddy Muñoz Sanabria. 2019. A videogame as a support tool for dyslexia therapy using the GDevelop platform. *ACM International Conference Proceeding Series* (2019), 162–166. https://doi.org/10.1145/3364138.3364172

[16] Nathan Gale, Pejman Mirza-Babaei, and Isabel Pedersen. 2015. Heuristic guidelines for wearable augmented reality applications. *CHI PLAY 2015 - Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (2015), 529–534. https://doi.org/10.1145/2793107.2810309

[17] Yan Ru Guo and Dion Hoe Lian Goh. 2016. Experimental Evaluation of Affective Embodied Agents in an Information Literacy Game. In *ICDL '16*. Newark, NJ, USA, 119–128. https://doi.org/10.1145/2910896.2910897

[18] Yan Ru Guo and Dion Hoe Lian Goh. 2016. From storyboard to software: User evaluation of an information literacy game. *Proceedings of the ACM Symposium on Applied Computing* 04-08-Apri (2016), 199–201. https://doi.org/10.1145/2851613.2851909

[19] Yan Ru Guo and Dion Hoe Lian Goh. 2016. Heuristic evaluation of an information literacy game. *Proceedings of the Association for Information Science and Technology* 53, 1 (2016), 1–4. https://doi.org/10.1002/pra2.2016.14505301107

[20] Celia Hodent. 2022. Advancing Game UX Maturity. *Game Usability* 2022 (2022), 7–24. https://doi.org/10.1201/9781003109389-3

[21] Kathryn Hymes, Jessica Hammer, Hakan Seyalioglu, Carol Dow-Richards, Deidra Brown, Trish Hambridge, Jill Ventrice, Meguey Baker, Yeonsoo Julian Kim, Tim Hutchings, and William S. Evans. 2021. Designing Game-Based Rehabilitation Experiences for People with Aphasia. *Proceedings of the ACM on Human-Computer Interaction* 5, CHIPLAY (2021). https://doi.org/10.1145/3474697

[22] Kyros Jalife, Casper Harteveld, and Christoffer Holmgård. 2021. From Flow to Fuse: A Cognitive Perspective. *Proceedings of the ACM on Human-Computer Interaction* 5, CHIPLAY (2021). https://doi.org/10.1145/3474683

[23] Kyros Jalife and Christoffer Holmgrd. 2019. Cognitive components of flow states and applications to video game design: A brief framework. *C and C 2019 - Proceedings of the 2019 Creativity and Cognition* (2019), 570–577. https://doi.org/10.1145/3325480.3326574

[24] Kalle Jegers. 2007. Pervasive game flow: Understanding player enjoyment in pervasive gaming. *ACM Computers in Entertainment (CIE)* 5, 1 (2007), 1–11. https://doi.org/10.1145/1236224.1236238

[25] Sumbul Khan, Attila Achenbach, Wei Quin Yow, and Lucienne Blessing. 2021. A Case Study on the Design of Touchscreen-Based User Interfaces for Multilingual Older Adults from Southeast Asian Backgrounds. *5th Asian CHI Symposium 2021* (2021), 167–173. https://doi.org/10.1145/3429360.3468204

[26] Kornchulee Khanana and Effie Lai Chong Law. 2013. Designing Children's Digital Games on Nutrition with Playability Heuristics. *Conference on Human Factors in Computing Systems - Proceedings* 2013-April (2013), 1071–1076. https://doi.org/10.1145/2468356.2468548

[27] Hannu Korhonen and Elina M.I. Koivisto. 2006. Playability heuristics for mobile games. *ACM International Conference Proceeding Series* 159, March (2006), 9–16. https://doi.org/10.1145/1152215.1152218

[28] Andrés Lucero, Jussi Holopainen, Elina Ollila, Riku Suomela, and Evangelos Karapanos. 2013. The Playful Experiences (PLEX) framework as a guide for expert evaluation. *Proceedings of the 6th International Conference on Designing Pleasurable Products and Interfaces, DPPI 2013* (2013), 221–230. https://doi.org/10.1145/2513506.2513530

[29] Rolf Molich and Jakob Nielsen. 1990. Improving a Human - Computer Dialogue. *Commun. ACM* 33, 3 (1990), 338–348.

[30] Christine Murad, Cosmin Munteanu, Benjamin R. Cowan, and Leigh Clark. 2021. Finding a New Voice: Transitioning Designers from GUI to VUI Design. In *CUI2021 - 3rd Conference on Conversational User Interfaces (CUI '21)*. ACM, New York, NY, USA, Bilbao (online), Spain, 12 pages. https://doi.org/10.1145/3469595.3469617

[31] Jakob Nielsen. 1992. Finding usability problems through heuristic evaluation. *Conference on Human Factors in Computing Systems - Proceedings* (1992), 373–380. https://doi.org/10.1145/142750.142834

[32] Jakob Nielsen. 1994. Enhancing the explanatory power of usability heuristics. *Conference on Human Factors in Computing Systems - Proceedings* (1994), 152–158. https://doi.org/10.1145/191666.191729

[33] Jakob Nielsen. 1994. Usability inspection methods. *Conference on Human Factors in Computing Systems - Proceedings* 1994-April (1994), 413–414. https://doi.org/10.1145/259963.260531

[34] Jukob Nielsen and Rolf Molich. 1990. Heuristic Evaluation Of User Interfaces. In *CHI 90 Proceedings*. 240–256.

[35] Zoë O'Shea and Jonathan Freeman. 2019. Game design frameworks: Where do we start? *ACM International Conference Proceeding Series* April (2019). https://doi.org/10.1145/3337722.3337753

[36] Janne Paavilainen. 2010. Critical review on video game evaluation heuristics. *Proceedings of the International Academic Conference on the Future of Game Design and Technology* (2010), 56–65. http://dl.acm.org/citation.cfm?id=1920787%5Cnhttp://portal.acm.org/citation.cfm?doid=1920778.1920787

[37] Janne Paavilainen. 2013. SoPlay Heuristics for Design and Evaluation of Social Network Games. *CHI 2013 Games User Research Workshop* January 2013 (2013), 6.

[38] Matthew J Page, Joanne E McKenzie, Patrick M Bossuyt, Isabelle Boutron, Tammy C Hoffmann, Cynthia D Mulrow, Larissa Shamseer, Jennifer M Tetzlaff, Elie A Akl, Sue E Brennan, Roger Chou, Julie Glanville, Jeremy M Grimshaw, Asbjørn Hróbjartsson, Manoj M Lalu, Tianjing Li, Elizabeth W Loder, Evan Mayo-Wilson, Steve McDonald, Luke A McGuinness, Lesley A Stewart, James Thomas, Andrea C Tricco, Vivian A Welch, Penny Whiting, and David Moher. 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *British Medical Journal* 372 (2021), 1–9. http://dx.doi.org/10.1136/bmj.n71

[39] Agnieszka Palalas and Norine Wark. 2017. Design Principles for an Adult Literacy Mobile Learning Solution. *ACM International Conference Proceeding Series* (2017). https://doi.org/10.1145/3136907.3136934

[40] Nathan Partlan, Erica Kleinman, Jim Howe, Sabbir Ahmad, Stacy Marsella, and Magy Seif El-nasr. 2021. Design-Driven Requirements for Computationally Co-Creative Game AI Design Tools. In *The 16th International Conference on the Foundations ofDigital Games (FDG'21)*. ACM, Montreal, QC, Canada, 12 pages. https://doi.org/10.1145/3472538.3472573

[41] Elaine Pearson and Adel Shaban. 2019. Evaluation of a prototype interactive working memory application for children with learning disabilities. *Proceedings of the 16th Web For All 2019 Personalization - Personalizing the Web, W4A 2019* (2019), 1–2. https://doi.org/10.1145/3315002.3332440

[42] David Pinelle, Nelson Wong, and Tadeusz Stach. 2008. Heuristic evaluation for games: Usability principles for video game design. *Conference on Human Factors in Computing Systems - Proceedings* (2008), 1453–1462. https://doi.org/10.1145/1357054.1357282

[43] David Pinelle, Nelson Wong, Tadeusz Stach, and Carl Gutwin. 2009. Usability heuristics for networked multiplayer games. *GROUP'09 - Proceedings of the 2009 ACM SIGCHI International Conference on Supporting Group Work* (2009), 169–178. https://doi.org/10.1145/1531674.1531700

[44] George Polya. 1945. *"Heuristic reasoning". How to Solve It: A New Aspect of Mathematical Method*. Princeton University Press, Princeton. https://doi.org/10.1515/9781400828678-039

[45] Aditya Ponnada and Ajaykumar Kannan. 2013. Evaluation of mobile games using playability heuristics. *Conference: Proceedings of the International Conference on Advances in Computing, Communications and Informatics* (2013), 244–247. https://doi.org/10.1007/978-3-319-02958-0{_}25

[46] Daniela Quiñones and Cristian Rusu. 2017. How to develop usability heuristics: A systematic literature review. *Computer Standards and Interfaces* 53, September 2016 (2017), 89–122. https://doi.org/10.1016/j.csi.2017.03.009

[47] Daniela Quiñones, Cristian Rusu, and Virginica Rusu. 2018. A methodology to develop usability/user experience heuristics. *Computer Standards and Interfaces* 59, November 2017 (2018), 109–129. https://doi.org/10.1016/j.csi.2018.03.002

[48] Florentin Rodio and J. M. Christian Bastien. 2013. Heuristics for Video Games Evaluation. , 89–93 pages. https://doi.org/10.1145/2534903.2534915

[49] Cristian Rusu, Silvana Roncagliolo, Virginica Rusu, and Cesar Collazos. 2011. A Methodology to Establish Usability Heuristics. *ACHI 2011 : The Fourth International Conference on Advances in Computer-Human Interactions* (2011), 59–62.

[50] Stephen R. Serge, Jonathan A. Stevens, and Latika Eifert. 2016. Make it usable: Highlighting the importance of improving the intuitiveness and usability of a computer-based training simulation. *Proceedings - Winter Simulation Conference* 2016-Febru (2016), 1056–1067. https://doi.org/10.1109/WSC.2015.7408233

[51] Chen Shu-Hui, Wu Wann-Yih, and Jason Dennison. 2018. Validation of EGameFlow: A self-report scale for measuring user experience in video game play. *Computers in Entertainment* 16, 3 (2018). https://doi.org/10.1145/3238249

[52] Gavin Sim. 2012. Designing the anti-heuristic game: A game which violates heuristics. In *IDC 2012, June 12–15*. ACM, Bremen, Germany, 308. https://doi.org/10.1145/2307096.2307153

[53] Gavin Sim, Janet C. Read, and Gilbert Cockton. 2009. Evidence based design of heuristics for computer assisted assessment. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 5726 LNCS, PART 1 (2009), 204–216. https://doi.org/10.1007/978-3-642-03655-2{_}25

[54] Maria Carmela Sogono and Deborah Richards. 2013. A design template for multisensory and multimodal games to train and test children for sound localisation acuity. *ACM International Conference Proceeding Series* (2013). https://doi.org/10.1145/2513002.2513005

[55] Penny Sweetser. 2020. GameFlow 2020: 15 Years of a Model of Player Enjoyment. *PervasiveHealth: Pervasive Computing Technologies for Healthcare* (2020), 705–711. https://doi.org/10.1145/3441000.3441048

[56] Penny Sweetser and Daniel Johnson. 2019. Gameflow and player experience measures: An initial comparison of conceptual constructs. *PervasiveHealth: Pervasive Computing Technologies for Healthcare* (2019), 317–321. https://doi.org/10.1145/3369457.3369486

[57] Penny Sweetser, Daniel Johnson, and Jay Kyburz. 2020. Evaluating GameFlow in a Multiplayer Online Strategy Game under Development. *PervasiveHealth: Pervasive Computing Technologies for Healthcare* (2020). https://doi.org/10.1145/3373017.3373068

[58] Penelope Sweetser, Daniel Johnson, and Peta Wyeth. 2012. Revisiting the GameFlow Model with Detailed Heuristics. *Journal of Creative Technologies* 2012, 3 (2012), 1–16. http://journal.colab.org.nz/article/21

[59] Penny Sweetser, Daniel Johnson, Peta Wyeth, and Anne Ozdowska. 2012. GameFlow Heuristics for Designing and Evaluating Real-Time Strategy Games. In *The 8th Australasian Conference on Interactive Entertainment*. The 8th Australasian Conference on Interactive Entertainment, Auckland, NZ.

[60] Penelope Sweetser, Daniel Johnson, Peta Wyeth, Aiman Anwar, Yan Meng, and Anne Ozdowska. 2017. GameFlow in different game genres and platforms. *Computers in Entertainment* 15, 3 (2017), 1–24. https://doi.org/10.1145/3034780

[61] Penny Sweetser and Zane Rogalewicz. 2020. Affording Enjoyment in VR Games: Possibilities, Pitfalls, and Perfection. *ACM International Conference Proceeding Series* (2020), 55–64. https://doi.org/10.1145/3441000.3441050

[62] Penny Sweetser, Zane Rogalewicz, and Qingyang Li. 2019. Understanding enjoyment in VR games with game flow. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST* (2019), 11–12. https://doi.org/10.1145/3359996.3364800

[63] Penelope Sweetser and Peta Wyeth. 2005. GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)* 3, 3 (2005). https://doi.org/10.1145/1077246.1077253

[64] Ramin Tadayon, Winslow Burleson, and Ashish Amresh. 2012. World of Golf: A socially relevant simulation game. *ACM International Conference Proceeding Series* (2012), 83–92. https://doi.org/10.1145/2367616.2367626

[65] Jean Lee Tan, Dion Hoe-Lian Goh, Rebecca P Ang, and Vivien S Huan. 2010. Usability and Playability Heuristics for Evaluation of an Instructional Game. *E-Learn 2010–World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* September 2016 (2010), 363–373. http://www.learntechlib.org/p/35570

[66] Gustavo F Tondello, Dennis Kappen, Marim Ganaba, and Lennart E Nacke. 2019. Gameful Design Heuristics: A Gamification Inspection Tool. In *Proceedings of HCI International*. 224–244.

[67] Gustavo F. Tondello, Dennis L. Kappen, Elisa D. Mekler, Marim Ganaba, and Lennart E. Nacke. 2016. Heuristic evaluation for gameful design. *CHI PLAY 2016 - Proceedings of the Annual Symposium on Computer-Human Interaction in Play Companion* (2016), 315–323. https://doi.org/10.1145/2968120.2987729

[68] Obelema Akobo Wodike, Gavin Sim, and Matthew Horton. 2014. Empowering teenagers to perform a heuristic evaluation of a game. *Proceedings of the 28th International BCS Human Computer Interaction Conference: Sand, Sea and Sky - Holiday HCI, HCI 2014* (2014), 353–358. https://doi.org/10.14236/ewic/hci2014.57

Received 2023-02-21; accepted 2023-07-07