EDeR

# Educational Design Research

Volume 8 | Issue 1 | 2024 | Article 61

Contribution Academic Article

Title **Research Slices: Core Processes for Effective Iteration in EDeR**

Author **Daniel Rees Lewis**
Northwestern University
USA

**Matthew Easterday**
Northwestern University
USA

**Chris Riesbeck**
Northwestern University
USA

Abstract Educational Design Research (EDeR) methodologists argue that iteration is a core component of EDeR (McKenney & Reeves, 2018). Iteration is currently defined as a process of gathering more information through actions, such as testing, and using that information to improve the design (Hoadley & Campos, 2022). In this paper, we seek to tighten the definition of iteration to help EDeR teams conduct iterations more effectively. We argue that EDeR teams should organize their research in *slices* that deliver small but real value to end users while informing the design research. EDeR should pick slices that are: (a) minimal and focused, (b) deployed in a real context, (c) valuable to the end users, and (d) informative to the research. Slicing helps EDeR teams increase ecological validity when they test because it allows testing which is within real-world educational contexts or with the stakeholders who will use and be impacted by the design. Increasing ecological validity of testing is particularly

Einige Autorinnen und Autoren dieses Special Issues sind Mitglieder des wissenschaftlichen Netzwerks Design-Based Research als methodologischer Rahmen in der Bildungsforschung (DBR-Netzwerk, gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – 452077361).

important because EDeR projects tackle highly complex real-world problems with many unknown elements and relational complexity – this means it is challenging to predict what designs will have the desired impact without real-world deployment. Effective iteration through organizing research in slices helps EDeR teams to better support stakeholder goals, develop more impactful theory, and have greater and earlier impact upon education.

**Keywords** Educational Design Research Methodology, Design Process, Iteration, Agile Research Processes, Lean, Slicing

**DOI** dx.doi.org/10.15460/eder.8.1.2127

**Citation** Rees Lewis, D., Easterday, M., Riesbeck, C. (2024). Research Slices: Core Processes for Effective Iteration in EdeR. *EDeR – Educational Design Research,* 8(1), 1-26.

dx.doi.org/10.15460/eder.8.1.2127

# Research Slices: Core Processes for Effective Iteration in EDeR

Daniel Rees Lewis, Matthew Easterday, Chris Riesbeck

## 1.0 Introduction

In this paper we argue for *slicing* in EDeR, a way of supporting iterating that creates and deploys useful parts of the design to provide immediate value to stakeholders and accelerate learning by the EDeR team. Slicing combats learning too late in EDeR projects. Slicing in research quickly discovers critical information about what design should be created by deploying designs in real-world settings.

EDeR teams seek to both meet stakeholder goals while producing novel theory (Collective, 2003; Collins et al., 2004; Easterday et al., 2016; Hoadley & Campos, 2022). EDeR teams "iterate toward better designs and add to foundational understanding of learning processes and how to support learning" (DiSalvo et al., 2017, p. 17). While EDeR researchers have stressed the importance of iteration to successful EDeR projects (Collective, 2003; Easterday et al., 2017; Gravemeijer & Cobb, 2006, 2013; Sannino et al., 2016), we in the EDeR research community have not defined specifically what good iteration looks like.

The current definition of iteration alone is not enough to ensure successful iteration. Iteration simply involves testing which in turn changes the EDeR team's understanding and changes something about the design (Adams et al., 2003; Schön, 1987) – EDeR teams might iterate successfully or unsuccessfully within these parameters. Thus, we must define what good iteration is, lest we create designs that stakeholders do not want (Penuel & Gallagher, 2017), or engage in an "elephantine effort result[ing] in the birth of mouse-like insights in their contribution to educational knowledge" (Dede, 2004, p. 107).

### 1.1 Effective Iteration is Required by Complexity and Ill-structuredness

EDeR teams attempt to create novel theories about how to create learning environments, including the theories about designs and learning processes (Collins et al., 2004; McKenney & Reeves, 2018). To do this, EDeR teams draw on existing theories about both designs and learning processes to inform what they create and help define what novel contribution they will make (Easterday et al., 2016; McKenney & Reeves, 2018). EDeR challenges are particularly complex because many aspects or variables of the problem impact each other in non-linear ways (relational complexity) (Brown, 1992; Collins et al., 2004;

Jonassen & Hung, 2015). EDeR teams need to consider student prior knowledge, teacher prior knowledge, learning goals, sequencing, relationships between students, relationships between student and teachers, relationships between stakeholders and the EDeR team, student and teacher reactions to designs, and much more. EDeR challenges are typically ill-structured, meaning that many of these aspects or variables of the problem that are not known cannot be predicted at the outset, and may change over time (Jonassen & Hung, 2015). For example, all the variables described above might all be somewhat unknown to start with and might all shift over the course of the EDeR project. Thus, complexity and ill-structuredness interact, as there are many aspects or variables to the problem which impact each other, with both many of the variables and their interactions being unknown.

Complexity and ill-structuredness of EDeR problems mean EDeR teams must iterate in order to understand enough to build working designs and novel theory. Iteration involves testing in some way – thought experiment, analysis, user test, or deployment – to generate a new understanding about the problem and then using that understanding to improve the design (Adams et al., 2003; Schön, 1987). EDeR teams iterate to co-evolve both their understanding of the problem, including learning processes, and the design itself (Cobb et al., 2003; Cross, 2011; Gravemeijer & Cobb, 2013). An EDeR team iterates out of necessity – if the team knew with certainty how to create an effective design they would simply create it. But the novel theory building challenges that they work on are rarely certain.

## 1.2 Related Research on Iteration in Design

The model presented here for iteration through slicing in design-based research is informed both by work in the design research community and the software product development community. The design community has focused on iterative learning. The software product development community has focused on the iterative delivery of end user value over very short timescales. Our contribution is to combine these two into an approach that combines rapid iterations of deployed value and learning.

In response to criticism that educational research does not address the complexity of learning environments (Hoadley, 2018; Lagemann, 2000; Vanderlinde & van Braak, 2010), EDeR researchers argued for conducting research in more realistic environments (Barab & Squire, 2004; Brown, 1992; Collins et al., 2004). Building off of participatory design perspectives (DiSalvo et al., 2017; Ehn, 1988), forms of EDeR such as formative evaluations stress rapid iterations involving those stakeholders responsible for carrying out the work of a given organization (Engeström, 2011). EDeR methodologists have noted that iteration should be cumulative – what is learned in one iteration influences the next iteration (micro-cycles; (Gravemeijer & Cobb, 2013; Gravemeijer & van Eerde, 2009). EDeR research is not a final step of quickly confirming results from the lab in real-world contexts.

Rather, theories of both learning processes and how to support those learning processes are developed through iterative EDeR (Barab & Squire, 2004; Brown, 1992).

EDeR is aligned with and directly influenced by approaches to design, such as design thinking, and human-centered design (Easterday et al., 2017; Hoadley, 2004). Such approaches focus on users and other stakeholders both when deriving the goals of the design and when testing to ensure that a design works as intended (Brown & Katz, 2011; Norman, 2013; Simon, 1996). Below we will outline examples that illustrate aspects of these approaches.

In the area of software product development, the Lean Startup (Ries, 2011) movement began in the early 2000s, inspired by the agile model for software development (Shore & Warden, 2021) that emphasized working in iterations of two weeks or less, rather than the traditional big releases of updated software every few years. Lean Startup used this idea to argue for very short build-measure-learn cycles for product development. This approach arose because of the high number of failures in product startups caused by the fact that most initial ideas turn out to be products that no one actually wants. Lean Startup argued for short build-measure-learn cycles on the order of weeks, not months. The goal of these iterations was not incremental implementation of a product idea, but learning what the product, if any, should be and for whom. In Lean Startup, as in design-based research, the focus is on testing hypotheses.

In a humorous sketch that went viral in the mid 2000's, Henrik Kniberg described what iterative rather than incremental development meant using the metaphor of developing a car (Figure 1; Kniberg, 2016). In the sketch, the incremental approach built a car from the bottom up, wheels then chassis then body and engine and eventually a working automobile. The iterative approach began by delivering a skateboard! Then a bicycle, a motorbike, and finally an automobile. There were two points being made: that iteration could and often did mean replacement of one idea with another, and that the second approach delivers value to the end user earlier.

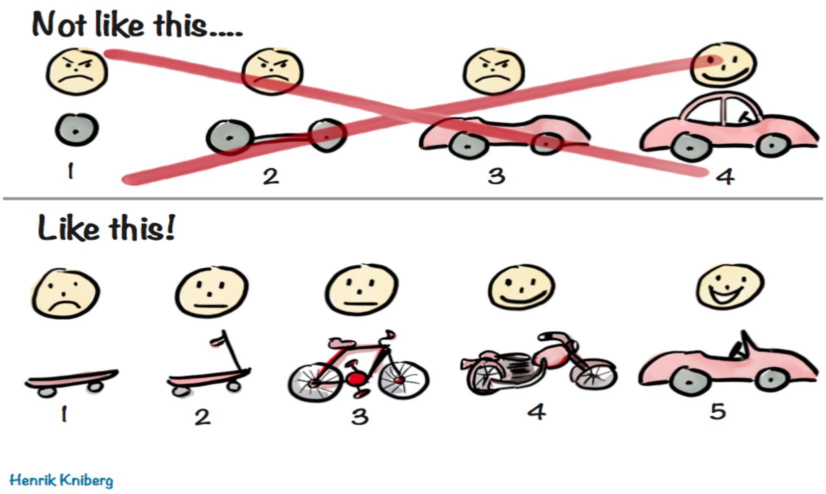


*Figure 1: Kniberg's sketch describing how iteration could involve replacement of ideas and should deliver value to the end user earlier.*

Kniberg's metaphor suggests one way to quickly deploy value: Use what already exists. When what is being designed is novel and needs to be built from scratch, a common agile technique with software applications is to do *vertical slicing* (Monday.com, 2022; Rasmusson, 2010). This is in contrast to *horizontal slicing* where an application is implemented by building out various layers: the database, the business logic code, the backend server to call that code, and the frontend web or mobile app to talk to the server. Vertical slicing means implementing just the bits of each layer needed to deliver one piece of testable functionality – e.g., just enough of user interface code, server code, business logic code, and data in a database to show a specific list of data to the user.

In recent years, researchers have sought to apply agile and lean concepts to their processes (Kirchherr, 2018) to help researchers collaborate and make more scientific progress. Labscrum is an agile approach that seeks to increase the amount of feedback within research labs in areas such as neuroscience, biology, and psychology through regular short meetings to early feedback on work in progress (May & Runyon, 2019). Developed a few years earlier in design research in human-computer interaction, Agile Research Studios employs a similar meeting structure and a set of support technology also to provide early feedback on work in progress of both research products and technology (Zhang et al., 2017).

Our work takes these themes of rapid iterative learning while delivering end user value as critical to effective research. Despite the widespread acknowledgement of the need for iteration in design, design processes provide relatively few principles for how EDeR teams pick the goal of the next iteration, a process we refer to as *slicing.*

## 1.3 Overview of Slicing for Iteration in EdeR

Slicing in general refers to the process of selecting one small element of the larger design problem for development and testing. We argue that slicing lies at the heart of effective iteration. There are many ways to slice badly, particularly when choosing the first slice. Take for example research on educational technology. Working on the login functionality for a novel educational app is very often a terrible first slice. While eventually that functionality must exist, building it would provide absolutely no opportunity for testing and learning about the viability or feasibility of the theoretical propositions implemented by the envisioned app. Furthermore, building and testing a login wouldn't be informative regarding either the extent the app supported learning processes or the learning processes themselves.

In what follows, we will give principles for choosing effective slices for EDeR with examples. Our slicing checklists posits that a slice should be: (a) **minimal**, (b) **deployed** in a real context, (c) **valuable** to the end users, and (d) **informative** to the research agenda. The remainder of this article will expand on these points. We begin with an illustrative example.

## 2.0 SLICING IN DETAIL

We illustrate slicing and non-slicing approaches through the fictional example project *Assist*, based on an actual multi-year research endeavor.

In project *Assist,* an EDeR team worked to support a K-12 teacher professional development training network. The teachers and teacher trainers in the network could only occasionally meet face-to-face, so much of the training occurred online. The network conducted professional development with hundreds of teachers across more than 20 school districts by training teachers to adapt their lesson to be more equitable and effective (Lewis et al., 2012). The network provided face-to-face feedback to teachers from teaching experts – university professors, and experienced teachers – to help improve their lessons. After negotiating with stakeholders (Penuel & Gallagher, 2017), the EDeR team scoped the following project: how to create technology that supported giving online feedback from teaching experts. Supporting more regular feedback to teachers emerged as the goal when the network was geographically distributed, creating challenges connecting experts and teachers.

The EDeR team sought to create an online platform to provide more frequent, timely, and higher quality feedback from these experts. Many experts were prepared to spend an hour a week giving feedback, but not multiple hours traveling to and from a training event.

We first present an approach to this project (1) *without* slicing followed by an approach (2) *with* slicing. In both examples the teachers pursued a learning and equity goal for improving their middle school math teaching. In both examples the EDeR team initially reasoned that a rubric-based feedback system would help provide organized feedback to teachers.

### 2.1 Non-Slicing Example

The EDeR team wanted to test if a rubric-based interface would provide feedback that would help teachers to improve their lessons. The EDeR team spent the five months developing a web application with pages to enable trainers to create rubrics, experts to give teachers feedback organized around those rubrics, and teachers to review that feedback. A notification system emailed teachers when feedback became available.

After these features were implemented, the EDeR team conducted three months of user testing. First, they showed the teachers sample feedback pages. The teachers made some cosmetic suggestions to change colors and rename buttons. In follow-up interviews, the teachers said that they felt that such feedback would have been useful.

The EDeR team then asked several experts to try using the new rubric-based system to provide feedback on samples of prior teacher work. The team observed that the notification feature sent an email every time any feedback was added, filling the teachers' inboxes with many emails; thus, the team changed the notification system to provide a summary email at the end of every day. The DBR team also saw that the experts had to constantly navigate between pages to review what they had written, because the rubric was split into sections on different webpages. So the EDeR team changed the interface so all the feedback is on one page.

Finally, after eight months of work, the EDeR team deployed the feedback system during live training. They asked experts to use the system to give feedback on a new lesson plan and teachers to review that feedback. For example, in the part of the rubric on supporting student thinking, one expert marked this area "needed work" and wrote: "The prompts in the first 10 minutes of class didn't always seem to encourage students to explain their thinking and how to work out area [of a shape]."

While the system functioned correctly and both experts and teachers were able to use it, the results were quite negative. The experts reported that they wouldn't use the software again because they felt like they were "grading" teachers. The teachers similarly reported they would reject the system for the same reason – they wanted to be helped to improve, not "told if we're good or not". Furthermore, the EDeR team could not find any evidence that the teachers used feedback given to improve their lessons. These emotional reactions were not revealed in user testing. They only became apparent when experts and teachers were engaged in actual coaching.

Despite following best practices, i.e., using a principled, user-centered design process, working iteratively, and producing a working prototype, after nine months the EDeR team was still in the very early stages of developing a serious design and model of coaching.

## 2.2 Slicing Example

Now consider how a slicing approach would have tackled the same situation.

The EDeR team wants to test if a rubric-based feedback system will help teachers to improve their lessons. They decide to first create and deploy a *concierge* version of the design, where the envisaged novel technological interaction is performed by the EDeR team (Courage & Baxter, 2005), using off-the-shelf technologies. This lets them quickly test basic features of the system with teachers and the experts.

The team spends several days working with the teacher trainers and experts to develop an initial list of rubrics for lesson plans. They create a Google spreadsheet with columns for each rubric to enter specific comments and scores ("excellent", "satisfactory", "needs work").

The EDeR team then deploys and tests the concierge system for two weeks with a teacher team in the middle of developing a lesson plan.

The experts enter their feedback into the rubric-structured spreadsheet, the EDeR team manually formats that feedback into a readable webpage, and sends an email to the teacher team with a link to that webpage to simulate a notification system.

The team observes both the experts providing feedback and the teacher team reviewing their feedback. On the plus side, the rubrics appear to encourage the experts to generate more feedback than previously seen. The process was less linear than expected. Experts jumped back and forth between columns, shifting and expanding comments under various rubrics, and changed the rubric scores multiple times.

In follow-up interviews, the EDeR team finds that both the experts and teachers reject the approach. The expert found scoring to be challenging. Both groups said the process felt too much like grading, and was inconsistent with how colleagues should interact. After just three weeks into the project, the team realizes that there is a serious issue with their envisaged design.

Since the rubrics seemed helpful for generating feedback, the EDeR team focuses on the scoring aspect. The team decides to do the simplest thing that could possibly work and drop scoring. The team eliminates the scoring columns from the spreadsheet and re-words the column headings to be questions of the form "do you have any comments on how well the lesson plan …" for each rubric. They adjust the web page report to look more like a series of chat comments, with section titles based on the rubrics.

The EDeR team then deploys and tests the new design for two weeks, using the concierge approach. The experts use the new spreadsheet to provide feedback on new work from the teacher team. The EDeR team collects the feedback into a web page and sends a notification email to the teacher team.

The same level of feedback is seen in the new system. Analysis of the feedback showed a greater mix of positive and negative elements within each rubric. For example, in response to a prompt about freedom of mathematical expression, an expert complimented the parts of the teaching team's lesson that had quality, and encouraged the teaching team to extend that aspect to the rest of their lesson. In the follow-up interviews, the experts said they felt less constrained to make their comments align with a score. The teaching team said they appreciated the breadth of feedback received. Neither group mentioned the grading issue, but the teaching team said that with the new feedback they weren't sure what the most important takeaways were, or what they should focus on next. They did, however, incorporate a suggestion to use math language routines to increase engagement into their next lesson plan.

The EDeR team moved into the next phase of their design and model of feedback. They saw that the givers of feedback wanted a system in which the input was conversational and less restrictive, but the receivers of feedback wanted output that was more organized and prioritized.

## 2.3 Comparison Between Examples

Let's compare the progress made in the two examples. The non-slicing example employed a process that reasonably follows many descriptions of design-based research and human-centered design. By the end of 9 months the team had gathered evidence that they had created a design that did not work. The team had not provided anything that the stakeholders themselves viewed as valuable. Compare this to the slicing example. By the end of 1 month the EDeR team had discounted the rubric approach, and by the end of 2.5 months the EDeR team had implemented a chat-stream approach that had helped teachers improve a lesson. The EDeR team had a design that, while not perfect, was able to function and that the stakeholders were willing to use. Critically, the EDeR team had created a feedback system that helped both the teachers improve their lesson and students to engage more in mathematical thinking in-class.

*Table 1: Slicing helps teams learn sooner by focusing on providing value to people*

| | **Non-slicing team** | **Slicing team** |
|---|---|---|
| **Learnings about the Design** | | |
| Single emails/per feedback leads to email deluge | Learned after 8 months | Learned after 1 month |
| Rubrics system made users felt like they were grading/being graded | Learned after 8 months | Learned after 1 month |
| Each piece of feedback might need multiple tags as knowledge; others connect ideas in feedback | Never | Learned after 2.5 months |
| Validated structured input system as a feedback system that helped teachers improve lessons | Never | Learned after 2.5 months |
| **Value for Stakeholders & Research Community** | | |
| Design concepts explored | Rubric approach | - Rubric approach<br>- Chat approach |
| Value provided to stakeholders | None | Improved lessons –> math language routines leading to more student engagement in lesson vs. baseline (2.5 months) |

In terms of learning, both teams learned that their rubric and email approaches did not work, and why. The slicing team made these discoveries after 1 month, while the non-slicing team made these discoveries after 8 months. Furthermore, the slicing team also learned how experts tended to write complex feedback that connected positive and negative feedback which required a different approach to their proposed tagging system. The slicing team also part-validated the chat-stream design – the design actually provided value in that it led to changes to a lesson which proved successful in increasing student in-class engagement. They did this all while also continuing to work on the design – they designed prompts, wrote email notifications, designed graphical interfaces, and coded working system functionality. Furthermore, the non-slicing team incorrectly believed they had validated the rubric approach early, but that was because their testing had elicited false-positive feedback – a pattern with testing that relies on user perception of a design that they do not use in a realistic context, something long noted by design scholars (Nielsen & Levy, 1994).

The important difference between the slicing and non-slicing examples is *not* about building more things more quickly, but instead the extent the EDeR team seeks to *build the right things for people earlier*. Some may misunderstand slicing as focusing on building a lot of things quickly – this is incorrect. Slicing as we define it focuses on providing value to the people the EDeR team serves. Slicing helps EDeR teams do less, not more, by focusing on testing theoretical assumptions with the minimal amount of effort. The slicing team spent far less time building than the non-slicing team and were closer to their goal of creating an effective feedback system and novel theory.

### 2.4 Definition of Slicing

We argue that EDeR teams should focus on picking effective slices when iterating, as illustrated in the slicing example above. Here, slicing is a process that involves building small, deployed, and useful parts of the design that aims to provide both value to stakeholders and learning to the EDeR team.

We can tighten this definition of what makes an effective slice. Specifically, a slice should be: (a) **minimal** and focused, (b) **deployed** in a real context, (c) **valuable** to the end users, and (d) **informative** to the research agenda. Core to slicing is the idea of stakeholder value – helping stakeholders achieve their goals. EDeR teams must also balance stakeholder goals with their own goals. For example, some stakeholder goals might run counter to the values of the EDeR team.

Implementing these principles can be challenging; EDeR teams typically work with busy and complex contexts. Nevertheless, we argue that teams should aim to follow these ideals as much as possible. Slicing is not always easy. It sometimes asks EDeR teams to do more things that they are less comfortable with – such as making relationships with stakeholders early and embedding themselves in a real learning environment with an early-stage design.

*Table 2: Properties of effective research slices*

| Attribute | Benefits | Challenges | Strategies |
|---|---|---|---|
| **A) Minimal**<br>The slice is as small as can be while delivering some specific value | Small slices are easier to make robust and deploy, less disruptive to end user, rest on fewer untested hypotheses, include fewer confounding changes | Untangling interdependencies in the implementation as well integration into the context of use; avoiding big design up front; avoiding over-building; delivering too little value | - Concierge testing and/or off-the-shelf solutions<br>- Do the simplest thing that could possibly work |
| **B) Deployed**<br>The slice is in real-world use by the target audience | Deployed use and monitor yields more ecologically valid results, earlier delivery of value, earlier testing of viability and feasibility | Reaching and getting buy-in from end users, developing a robust useful deliverable, integrating into existing information and action flows; overcoming fear of commitment and/or failure | - Concierge testing and/or off-the-shelf solutions<br>- Parallel system<br>- Early adopters<br>- Watchful deployment<br>- Explicit risk-benefit cost of delay analysis |
| **C) Valuable**<br>The slice provides significant value to the end users | Delivering value leads to greater ecologically valid use, creative and informative adaptation by end users, more ethical, less extractive research | Identifying actual user value and pain points, supporting learning and usage during and after the research period | - Do the simplest thing that could possibly work<br>- Human-centered design research/ participatory design |
| **D) Informative**<br>The slice is designed to significantly advance the research agenda | Proper experimental design and monitoring leads to less risk of failure, greater information gain, collection of broader data related to social issues and stakeholder impact beyond end user value | Identifying which hypotheses in theoretical model to explore first, aligning those questions with user value, designing methods of data collection with low user friction, finding short-term proxies when long-term impacts are involved | - Explicit risk-benefit cost of delay analysis<br>- Watchful deployment |

**A) Minimal**. The slice is as small as can be while delivering some specific value to stakeholders. Minimal slices are beneficial because they are easier to make robust and deploy, thus avoiding wasted work of building a big failed design, and are less disruptive to end users. Furthermore, by being small the design rests on fewer untested hypotheses and its introduction includes fewer potential confounding changes – that is, because fewer things are changed, EDeR teams can more easily untangle the causal effects of any changes observed.

Take for example how the slicing team was able to deploy their design within a few weeks. They worked out the minimal thing they would need to build in order to provide potentially valuable feedback to teachers. Conversely, the non-slicing team spent five months building a functional design that did not work.

**B) Deployed**. The slice is in the real world used by the target audience as part of their actual practice and not part of a hypothetical scenario. Deploying and monitoring is beneficial because it yields more ecologically valid results (Kihlstrom, 2021) and earlier delivery of value to stakeholders. Furthermore, it produces earlier testing of viability and feasibility within real-world contexts.

Take our example above where the EDeR team that used the slicing approach gave their designs to the actual stakeholders to use – teachers and experts. Furthermore, the team gave feedback on a lesson as the teachers were working on improving it; unlike the non-slicing approach, the feedback was not hypothetical so the users were not being asked to react to a hypothetical situation.

**C) Valuable.** The slice provides significant value to the end users – or at least represents the EDeR team's genuine efforts to do so. Delivering value is beneficial as it leads to greater ecologically valid use, as users are reacting to things they actually care about. Furthermore, the EDeR team might see creative and informative adaptation by end users that can both inform them of user goals and ways to improve the design (Barab & Squire, 2004). Finally, doing so offers the potential to lead to more ethical, less extractive research.

Take for example how the slicing team provided value for the teachers, as they were able to provide feedback that the teachers used to improve their lessons. Conversely, the non-slicing team did not produce value for the teachers or the teaching experts in their first deployment. This was because the non-slicing team did not create a system that provided feedback on a lesson that teachers were working on.

**D) Informative.** The slice is designed to significantly advance the research agenda, such as better understanding learning processes, or how to support learning processes, or both. Being informative is beneficial as proper experimental design and monitoring leads to less risk of failure and gaining more information. Furthermore, it allows for the collection of broader data related to social issues and stakeholder impact beyond end user value.

Take for example how the slicing team were able to move their theory of how to support learning processes away from a failed approach – a rubric-based system – towards a better theory.

In section 3 below we unpack the challenges and strategies noted in Table 2.

## 2.5 Sketch of Iteration for EDeR

We have adapted Kniberg's metaphorical sketch of product development (Figure 1; Kniberg, 2016) to incorporate both users and the EDeR team (Figure 2). Where Kniberg focused on the delivery of value to users, we add the delivery of insight to the researchers. Where Kniberg assumed an automobile was the ultimate known goal, we assume the fog of innovation where both EDeR teams are exploring what might work. We use a penny-farthing with the non-slicing example to represent the building of a solution that does not work – as occurred with the non-slicing EDeR team in section 2.1.

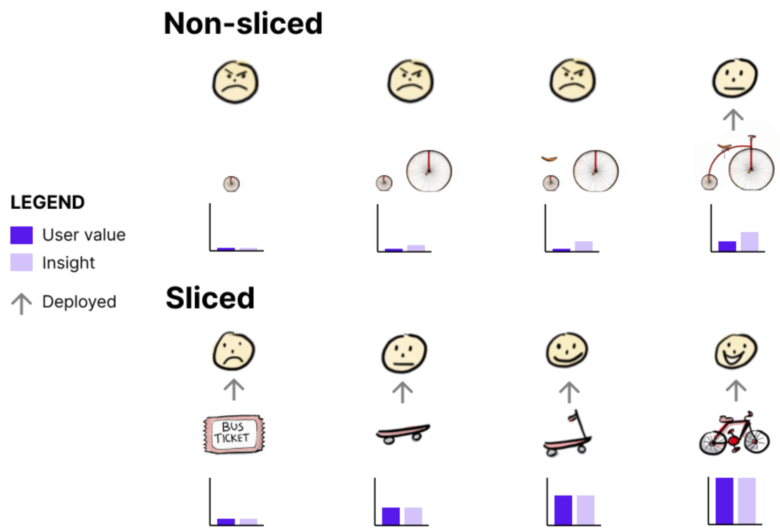


*Figure 2: Adaptation of Kniberg's sketch for EDeR to demonstrate how slicing can more quickly deliver initial insights and user value and avoid building the wrong thing.*

As in Kniberg's sketch, the non-slicing team delivers no value to users for many weeks. The team does gain some insights about their envisioned product during construction and testing but learn nothing about how that product might perform with real users in the real world. In contrast, almost immediately the slicing team deploys a potential solution – bus tickets – and learns that their users, while valuing the relatively low cost of mass transit, find schedules and available endpoints too constraining. They learn that users need a solution that is available beyond commercial business hours.

While the non-slicing team is still iterating internally on the materials and mechanics of a penny-farthing, the slicing team engages in several experiments with cheap, simple human-powered devices. They learn that users value steerability, but would trade small size for greater speed and the ability to travel on streets rather than sidewalks.

At the point where both teams deliver their first prototype, the relationship between researchers and users has diverged dramatically. In the non-slicing scenario, the penny-farthing is the first thing the users have seen from the research team in months. While understanding that this is a prototype, it is hard not to feel that some of the issues with the solution, such as mounting the penny-farthing and parking it, should have been obvious if only the researchers had spent more time with the users and their travel needs. The users in the slicing scenario, on the other hand, have been in frequent contact with the EDeR team. They have been part of the design process. The prototype bicycle is no surprise, but something raised and agreed upon a short time before.

## 2.6 The Overall Benefits of Slicing

Slicing as defined here can lead to more accurate learning because it has high ecological validity. The EDeR team will have greater confidence in what is learned about the impact of their current design because the design has been deployed in real contexts. This learning includes surfacing problems and design considerations the team hasn't considered – unknown unknowns (Bammer, 2008). Deployment forces the design team to face the complexity of the real world – putting the theory "in harm's way" (Cobb et al., 2003, p. 10).

Iteration with effective slicing allows the EDeR team to reduce wasted work. It is all too easy to litter the world with unrealized designs that teams have spent a significant amount of building that stakeholders do not want and that do not advance theory (Dede, 2004; Kniberg, 2016). The non-slicing team in the example above spent 8 months building the rubric system only to find out that it provided no value to stakeholders and no advancement of theory as to what types of designs might work in this context.

## 2.7 Design-based Ecological Validity

In this work, we emphasize the concept of design-based ecological validity. In EDeR ecological validity is associated with the extent the design is deployed into the real context that is being designed for, with real stakes for the stakeholders (Barab & Squire, 2004). Psychologists use ecological validity to mean either (a) how similar the lab conditions are to contexts outside of the lab or (b) how much the findings apply to contexts outside of the lab (Kihlstrom, 2021). The latter is how some EDeR scholars define ecological validity – focusing on the outcome of the findings (e.g. Euler, 2014; Gravemeijer & Cobb, 2013). We use the former definition, the original use of the term defined by Brunswik (1956).

Just as with the definition of traditional ecological validity (Kihlstrom, 2021), design-based ecological validity is not a single dimension or variable. For example, the test of the rubric system by the non-slice-based EDeR team had some ecological validity in terms of using real stakeholders and using real lessons. However, it was much less ecologically valid in terms of timing within the teacher training process – the teachers had not just developed the lesson and were not able to apply the feedback. Because the test was low on this variable of ecological validity, the team did not learn vital information about their rubric system until they deployed the system later. Our use of the term deployment seeks to maximize ecological validity across variables, as it is the real context of use.

## 2.8 The Ethics of Slicing and Iteration

In this section we want to both acknowledge that slicing and iteration must address ethical issues and show how slicing might support an EDeR team's ethical commitments. Deployment necessarily requires significant ethical considerations.

**Cost of action:** Deploying means there are real stakes for stakeholders – we must be careful that an ineffective design does not make things worse, even for a short period of time. Critics of a "move fast and break things" attitude have noted that designers can do harm in their desire to learn quickly (Vardi, 2018). We reiterate that slicing is not about moving more quickly and doing more things – rather, it is about building small, with end user value as a primary criterion, and testing early. Slicing is about *watchful deployment* – EDeR teams are present for deployments in which they can halt and rectify any potential harm caused early.

**Cost of inaction:** In any ethical analysis we should also include any ongoing harms that are occurring in the status quo. That is, there is a cost to not acting. Take the context described in the example above – in the status quo, the education system in California is leading to much lower test scores for English language learners. Lower math test scores in turn have negative socio-economic impacts on those students (Chetty et al., 2018).

Our claim here is not that slicing will necessarily lead to more ethical research practices and outcomes. Slicing is a methodological tool, and like all tools it can have a range of impacts from an ethical perspective. Rather, we claim that slicing can offer more opportunities for more ethical practices – because of its focus on providing value to stakeholders, giving stakeholders opportunities to understand and reject the design, and emphasizing the importance of developing slices that inform the larger research questions, beyond specific end user value.

**Extraction:** There have been criticisms, rightly in our view, of EDeR and other educational research as being extractive (Vanderlinde & van Braak, 2010). Researchers enter a community, gather data, and then leave to write their papers, having provided little if any value to the other stakeholders. By having a primary focus on deploying value to stakeholders early, EDeR teams might better produce both value to stakeholders and develop more practical theory.

**Power focus:** In alignment with other educational scholars, we argue that a core component to the ethics of education research is power (Hand et al., 2013; Vakil et al., 2016). Slicing can support more frequent opportunities for stakeholder use of power. This is because they can have a more accurate understanding from which to give that feedback. It is easy for stakeholders to say yes to a design and plans when they are abstract only to realize later that the design does not provide value, or worse, causes harm. Regular deployment can give more opportunities for stakeholder power from a position of more knowledge.

## 3.0 CASE STUDIES IN SLICING PITFALLS, CHALLENGES AND STRATEGIES

Slicing can be challenging. It can seem impossible to do and attempts can fail. To illustrate common obstacles and pitfalls, we present four different case studies on how slicing is challenging in EDeR projects and some relevant strategies to use in developing slices.

### 3.1 Pitfall: Unbounded Internal Review

In this case study, a design project spent two years in the design phase, with no deployment of value, even though there were no major obstacles, technical, conceptual, or otherwise.

An EDeR team of education and medical researchers were working on a project to help people recently diagnosed with a specific chronic illness to manage that illness. The need was clear: 30% of patients would be hospitalized within 1 month of diagnosis. Self-management for this illness involved significant changes in diet and exercise, daily monitoring of metabolic levels. Compliance failure was a major factor in hospitalization. The EDeR team's task was to develop a complete set of online and paper-based self-care instructions for patients to use at home. This phase took a little over six months.

Because patients would be using these instructions on their own at home with no in-hospital training, the usability of the materials was of great concern. Consequently, the EDeR team conducted multiple rounds of usability testing for eight months, with proxy users with similar age and educational profiles as the patients, until no critical usability issues arose, even with older adults with limited experience with technology or medical language. The team then conducted five more months of multiple rounds of expert review of the instructional materials with patient educators and doctors. Because this led to changes in the instructions, the EDeR team conducted another five months of usability testing.

After two years, the project had not delivered any value to the patient population. The team still did not know how actual patients in actual home environments would respond to their design. This happened despite a well-understood development plan, validated technical content, careful design guided by multimedia learning theory (Mayer, 2002), and spiral learning (Harden, 1999), ready access to a target user group, and clear metrics for success based on patient compliance with diet and medication management and the rate of hospitalization.

Our analysis suggests that fear of deployment within the EDeR team was a major factor. This fear arose from at least two sources: the justifiable fear in any medical design of doing harm, and the fear in academic research of expending great effort and resources to no discernible or publishable effect. These fears led to a cycle of endless refinement and delay. The perfect became the enemy of the good.

Two strategies for slicing and deploying much earlier and more frequently could have addressed these fears:

- **Explicit risk/benefit/cost of delay analysis**: The medical EDeR team spent most of their effort mitigating just one risk: *Usability*. Risk analysis (Carlson et al., 2020) would have raised other risks, including the *feasibility* of delivering instructions to patients in the hospital system, the *viability* of the instructions when used in a home context, and the *desirability* to the patients of attending to the instructions when sick and stressed at home. Cost of delay analysis (Reinertsen, 2009) would have quantified the cost to patients of delaying deployment. This analysis would have argued strongly for more ecologically valid testing through early deployment to test multiple risks, not just usability.

- **Watchful deployment**: It is a misconception to view deployment in slicing as simply releasing a design into the wild. Deployment in slicing is a form of ecologically valid testing that requires careful monitoring and logging in order to identify if value is actually being delivered and, if not, why not. Deploying small changes reduces the potential for harm, causes less disruption for patients, and reduces the number of confounding factors in the observations.

## 3.2 Pitfall: Building Too Much

A common pitfall is building too much, too soon, with too many unvalidated assumptions.

An EDeR team was working on a project to provide just-in-time case studies of teaching activities to trainers coaching K-12 teachers in a state-wide learning network. The team decided to develop a web-based real-time collaborative case study authoring system to allow teachers to write examples of their personal lesson plans that trainers could share with other teachers. After more than a year of design and development, the end result was an authoring tool with bugs and no users.

Why did the project fail? As with many software projects (Little, 2006), the team drastically underestimated how hard it would be to implement the many features in their design and still have a simple, easy-to-use interface. They overestimated the level of effort teachers would be willing to invest in writing lesson plan case studies.

Had the team focused on early deployment of high-value informative slices, they could have had greater impact and learning in a much shorter period of time. For example, they could have begun by organizing workshops with teachers and trainers to share experiences about successful and unsuccessful lesson plans. This non-technical low-commitment request would have revealed the quantity and quality of such experiences and the level of interest in sharing them. If successful, the team could then have documented and shared those lesson plans with other trainers and teachers to explore the content and structure of lesson plan experiences, how valuable teachers found these plans, and what other aspects of lesson plans teachers most wanted to see. If things still looked promising, the team would then be ready to design and test a minimal lesson plan case study editor with the teachers who participated most in the earlier meetings. In this way, each step would have delivered value or provided important information on flaws in the core idea.

Here are a few common strategies to avoid the pitfall of building too much too soon:

- **Concierge testing**: Start by delivering value manually. Delay automation until value is demonstrated in an ecologically valid way as possible, and specific needs are clear and well-defined.
- **Do the simplest thing that could possibly work:** Scale up using common off-the-shelf tools, such as email, shared spreadsheets, and web forms. These tools are already familiar to users. Let a project scale up past the concierge point, and provide information about what functionalities might be needed when developing a new tool, if any.

## 3.3 Challenge: Slicing Complex Interdependencies

Deployment of slices of value can seem impossible when the change that is planned disrupts many interdependent parts and stakeholders of a complex system. This occurs often in organizational change.

For example, the faculty at a school for education and social policy wanted to revamp their undergraduate civic certificate program to provide students with practical hands-on experience in community organization and change, while delivering actual value to the community and city government. The existing program had a two-quarter capstone where students only wrote but did not implement a set of recommendations. The vision was to replace this with a two-year sequence working with community partners to develop proposals, plan community events, and implement a city-wide get-out-the-vote campaign.

The challenge that the faculty faced was how to make a major multi-year change to a curriculum without disrupting student progress, overloading the faculty with double duty, or confusing existing community partner relationships.

The faculty sliced this complex change by first introducing a new optional first-year “cornerstone” project course taught by a new faculty. This allowed that faculty to pilot the new curriculum with early adopters. In the following year, all of the new first-year students took the piloted cornerstone, while second-year students had the option of continuing in the cornerstone course, or taking the old capstone course. In the third year, the capstone was retired.

This phased approach delivered value early to students, faculty, and the community. The students who wanted it were given an immediate opportunity to gain real-world experience in community organizing, but students who had planned on the existing curriculum were not disrupted. The faculty who was planning the new curriculum was able to run a pilot offering of the cornerstone course with early adopters. The pilot exposed a number of issues with project-based learning in this context without the confounding issues of mixed cohorts. The community partners were eased into a new model of working with students over a multi-year timespan that was compatible with the pace of change in the local government.

Two strategies were used here to deal with making large changes to a complex system with many interdependencies:

- **Set up a parallel system:** Instead of replacing the two-quarter capstone class with a two-year project class, parts of the new curriculum were introduced in parallel. The old curriculum was retired only when it was no longer needed. This is analogous to how legacy corporate computer systems are often replaced. A new system is set up, but initially it just handles one functionality, typically something the old system does not do or does poorly. Gradually, and invisibly to end users, more functions are migrated to the new system, until eventually the legacy system can be retired. This approach allows ample opportunity for discovering and fixing deficiencies in the new system.
- **Start with early adopters:** To maximize the ratio of gain to pain, early slices are deployed to users most in need of the delivered value, and/or least affected by the changes being made (Moore & McKenna, 1999; Rogers, 2010).

## 4.0 SUMMARY

In this article, we have presented four characteristics of slices for effective iteration in design-based research: minimal and focused, deployed, valuable to end users, and informative to the research goals. These four characteristics are necessary and interdependent:

- Slices need to be minimal and focused to support frequent deployments, earlier delivery of value, reduced disruption to stakeholders, fewer confounding factors for research, and more rapid learning.

- Slices need to be deployed to deliver value and provide ecological validity to the research learning.

- Slices need to deliver value to end users to ensure ecologically valid use and generate informative research insight by identifying potential additional benefits and areas of friction.

- Slices need to integrate with the research program to maximize learning, provide a coherent path for future iterations, and take into account larger social and ethical issues beyond product goals of functionality and usability.

Achieving all four criteria can be challenging. For example, EDeR teams may be unable to deploy in realistic environments. The IBM Watson project to develop a system to win at Jeopardy could not be tested against Jeopardy champions on a regular basis. Their solution was to analyze 2000 Jeopardy games to develop a formal model of champion-level behavior, against which they could repeatedly evaluate their evolving prototype (Ferrucci et al., 2010). For another example, one author worked on a medical educational technology project to teach patients recently diagnosed with a chronic disease how to manage their illness. For legal reasons, early designs could not be tested with real patients in this situation, but more experienced patients knew too much to be realistic participants. Therefore, the EDeR team tested the ability of healthy subjects in the same age group who were unfamiliar with the disease to follow the instructions in early prototypes. This allowed the team to iterate instructional content until the risks of miscommunication were minimal. It remains important to limit the number of such non-deployment iterations because they deliver no value to real patients and lack ecological validity.

Another issue is deciding what a slice should focus on. Typically, there are a number of different values to users that could be deployed in a project. Which one would be most informative for the research program? To answer this question, we argue elsewhere for applying the concepts of *causal chains* and *risking* (Carlson et al., 2018). EDeR teams should pick slices that address the biggest risks (uncertainty) in their causal theory of change (Carlson et al., 2020). For example, in the Assist project slicing example, the teams' success depended on early testing of the risky assumption that the rubric would improve feedback from teaching experts.

Addressing risks in causal chains is especially tricky in education where we seek to intervene on long causal chains. For example, in the *Assist* project, the EDeR team's causal theory of change assumed that a rubric would lead to better feedback from teaching experts, to changes to the lesson plan that might improve teacher learning directly, and also to changes in the lesson that would lead to changes in student behavior, then to changes in student learning, creating additional teacher learning that would influence future lessons (and so on). Longer term outcomes such as transfer and educational attainment or life outcomes might seem impossible to address with a short-term slice.

In order to pick slices most likely to produce value and research learnings, the EDeR must first explicitly identify their causal assumptions and which assumptions are most uncertain. For example, based on current educational theory, one can reasonably bet that improving teacher learning will lead to better student outcomes, so the team can work backward from a longer-term outcome (student learning) by using proxy (teacher learning) that they believe will causally influence the outcome. The Assist project teams believed their biggest risk was whether the rubric would produce better feedback from teaching experts and started working forward from the first step in this causal change. In the slice example, the EDeR team was able to focus on testing their riskiest causal assumption using a simple mockup, which then led to greater impact on the downstream, less risky parts of the chain.

There is still much to explore in the area of slices for iterated EDeR. The catalog of strategies to deal with common obstacles and pitfalls to slicing needs to be extended. A taxonomy of challenges and which strategies are most useful for dealing with them could then be developed.

## 5.0 Acknowledgment

We thank Elizabeth Gerber, Spencer Carlson, Haoqi Zhang, Kristine Lu, Gus Umbelino, Evey Huang, and the Delta Lab for helping us talk through these ideas over many years. Furthermore, we thank the organizers and attendees of 2022 Symposium Knowledge by Design in Education: Getting to the Core of DBR Epistemology whose questions and discussions on our presentation helped us sharpen these ideas. Finally, thank you to the editors and reviewers whose thoughtful feedback greatly improved this paper.

This work was supported through funding by the National Science Foundation (IIS-1530833 and IIS-2008450), and the Spencer Postdoctoral Research Fellowship received by the first author. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or other funding bodies.

This work greatly extends an earlier and less well-developed conference paper presented at the International Conference of the Learning Sciences (Rees Lewis et al., 2020).

## 6.0 References


Adams, R. S., Turns, J., & Atman, C. J. (2003). Educating effective engineering designers: The role of reflective practice. *Design Studies*, *24*(3), 275–294. https://doi.org/10.1016/S0142-694X(02)00056-X

Bammer, G. (2008). Do we need a new discipline to document and transmit problem-based learnings? *Proceedings of the 11th SDPS Transdisciplinary Conference on Integrated Systems, Design, & Process* Science *at Asia University Taichung, Taiwan, June 1-6, 2008 (IDPT-2008).*

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*, *13*(1), 1–14.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, *2*(2), 141–178.

Brown, T., & Katz, B. (2011). Change by design. *Journal of Product Innovation Management*, *28*(3), 381–383.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. University of California Press.

Carlson, S. E., Maliakal, L. V., Rees Lewis, D. G., Gorson, J., Gerber, E., & Easterday, M. (2018). Defining and assessing risk analysis: The key to strategic iteration in real-world problem solving. *Rethinking Learning in the Digital Age: Making the Learning Sciences Count*, 1, 352–359.

Carlson, S. E., Rees Lewis, D. G., Maliakal, L. V., Gerber, E. M., & Easterday, M. W. (2020). The design risks framework: Understanding metacognition for iteration. *Design Studies*, *70*, 100961. https://doi.org/10.1016/j.destud.2020.100961

Chetty, R., Friedman, J. N., Hendren, N., Jones, M. R., & Porter, S. R. (2018). *The opportunity atlas: Mapping the childhood roots of social mobility*. National Bureau of Economic Research.

Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9–13.

Collective, D.-B. R. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, *32*(1), 5–8.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, *13*(1), 15–42.

Courage, C., & Baxter, K. (2005). *Understanding your users: A practical guide to user requirements methods, tools, and techniques*. Gulf Professional Publishing.

Cross, N. (2011). *Design thinking: Understanding how designers think and work*. Berg.

Dede, C. (2004). If design-based research is the answer, what is the question? A commentary on Collins, Joseph, and Bielaczyc; diSessa and Cobb; and Fishman, Marx, Blumenthal, Krajcik, and Soloway in the JLS special issue on design-based research. *The Journal of the Learning Sciences*, *13*(1), 105–114.

DiSalvo, B., Yip, J., Bonsignore, E., & Carl, D. (2017). Participatory design for learning. In *Participatory Design for Learning* (pp. 3–6). Routledge.

Easterday, M. W., Rees Lewis, D. G., & Gerber, E. M. (2017). The logic of design research. *Learning: Research and Practice*, 1–30. https://doi.org/10.1080/23735082.2017.1286367

Easterday, M. W., Rees Lewis, D. G., & Gerber, E. M. (2016). The logic of the theoretical and practical products of design research. *Australasian Journal of Educational Technology*, *32*(4).

Ehn, P. (1988). *Work-oriented design of computer artifacts* [PhD Thesis]. Arbetslivscentrum.

Engeström, Y. (2011). From design experiments to formative interventions. *Theory & Psychology*, *21*(5), 598–628.

Euler, D. (2014). Design-Research – a paradigm under development. In D. Euler & P. F. E. Sloane (Eds.), *Design-based research. 27. Beiheft der Zeitschrift für Berufs- und Wirtschaftspädagogik (ZBW)* (pp. 15–44). Franz Steiner Verlag.

Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A. A., Lally, A., Murdock, J. W., Nyberg, E., & Prager, J. (2010). Building Watson: An overview of the DeepQA project. *AI Magazine*, *31*(3), 59–79.

Gravemeijer, K., & Cobb, P. (2013). Design research from a learning design perspective. In T. Plomp & N. Nieveen (Eds.), *Educational design research* (pp. 72–113). Netherlands Institute for Curriculum Development (SLO).

Gravemeijer, K., & Cobb, P. (2006). Design research from a learning design perspective. In *Educational design research* (pp. 29–63). Routledge.

Gravemeijer, K., & van Eerde, D. (2009). Design research as a means for building a knowledge base for teachers and teaching in mathematics education. *The Elementary School Journal*, *109*(5), 510–524.

Hand, V., Penuel, W. R., & Gutiérrez, K. D. (2013). (Re)framing educational possibility: Attending to power and equity in shaping access to and within learning opportunities. *Human Development*, *55*(5–6), 250–268.

Harden, R. M. (1999). What is a spiral curriculum? *Medical Teacher*, *21*(2), 141–143.

Hoadley, C. (2018). A short history of the learning sciences. In *International Handbook of the Learning Sciences* (pp. 11–23). Routledge.

Hoadley, C. M. (2004). Learning and design: Why the learning sciences and instructional systems need each other. *Educational Technology, 44(3)*, 6–12.

Hoadley, C., & Campos, F. C. (2022). Design-based research: What it is and why it matters to studying online learning. *Educational Psychologist*, *57*(3), 207–220.

Jonassen, D. H., & Hung, W. (2015). All problems are not equal: Implications for problem-based learning. In A. Walker, H. Leary, C. Hmelo-Silver, & P. A. Ertmer (Eds.), *Essential Readings in Problem-Based Learning* (pp. 7–41). Purdue University Press.

Kihlstrom, J. F. (2021). Ecological validity and “ecological validity.” *Perspectives on Psychological Science*, *16*(2), 466–471.

Kirchherr, J. (2018). *The lean PhD: Radically improve the efficiency, quality and impact of your research*. Bloomsbury Publishing.

Kniberg, H. (2016, January 25). Making sense of MVP (Minimum Viable Product) – and why I prefer Earliest Testable/Usable/Lovable. *Crisp's Blog*. https://blog.crisp.se/2016/01/25/henrikkniberg/making-sense-of-mvp. Accessed 18 July 2023.

Lagemann, E. C. (2000). *An elusive science: The troubling history of education research*. University of Chicago Press.

Lewis, C. C., Perry, R. R., Friedkin, S., & Roth, J. R. (2012). Improving teaching does improve teachers: Evidence from lesson study. *Journal of Teacher Education*, *63*(5), 368–375.

Little, T. (2006). Schedule estimation and uncertainty surrounding the cone of uncertainty. *IEEE Software*, *23*(3), 48–54.

May, L., & Runyon, T. (2019). *LabScrum: A case study for agility in academic research labs*. https://doi.org/10.31234/osf.io/zg4ub

Mayer, R. E. (2002). Multimedia learning. In *Psychology of learning and motivation* (Vol. 41, pp. 85–139). Elsevier.

McKenney, S., & Reeves, T. C. (2018). *Conducting educational design research*. Routledge.

*Monday.com* (2022, September 9). What vertical slicing is and how it's used in project management. https://monday.com/blog/project-management/vertical-slice/. Accessed 20 July 2023.

Moore, G. A., & McKenna, R. (1999). *Crossing the chasm: Marketing and Selling High-tech Products to Mainstream Customers*. Harper Business.

Nielsen, J., & Levy, J. (1994). Measuring usability: Preference vs. performance. *Communications of the ACM*, *37*(4), 66–75.

Norman, D. A. (2013). *The design of everyday things* (Revised and expanded edition). Basic Books.

Penuel, W. R., & Gallagher, D. J. (2017). *Creating Research Practice Partnerships in Education.* ERIC.

Rasmusson, J. (2010). *The agile samurai: How agile masters deliver great software*. The Pragmatic Bookshelf.

Rees Lewis, D. G., Carlson, S. E., Riesbeck, C. K., Lu, K. J., Gerber, E. M., & Easterday, M. W. (2020). *The logic of effective iteration in design-based research.* Paper presented at the Proceedings of the 14th International Conference of the Learning Sciences, Nashville, TN.

Reinertsen, D. G. (2009). *The principles of product development flow: second generation lean product development*. Celeritas Publishing.

Ries, E. (2011). *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses*. Crown Business.

Rogers, E. M. (2010). *Diffusion of innovations*. Simon and Schuster.

Sannino, A., Engeström, Y., & Lemos, M. (2016). Formative interventions for expansive learning and transformative agency. *Journal of the Learning Sciences*, *25*(4), 599–633.

Schön, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions* (pp. xvii, 355). Jossey-Bass.

Shore, J., & Warden, S. (2021). *The art of agile development*. O'Reilly Media, Inc.

Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). MIT Press.

Vakil, S., McKinney de Royston, M., Suad Nasir, N., & Kirshner, B. (2016). Rethinking race and power in design-based research: Reflections from the field. *Cognition and Instruction*, *34*(3), 194–209.

Vanderlinde, R., & van Braak, J. (2010). The gap between educational research and practice: Views of teachers, school leaders, intermediaries and researchers. *British Educational Research Journal*, *36*(2), 299–316.

Vardi, M. Y. (2018). Move fast and break things. In *Communications of the ACM* (Vol. 61, Issue 9, p. 7). ACM New York.

Zhang, H., Easterday, M. W., Gerber, E. M., Rees Lewis, D., & Maliakal, L. (2017). Agile research studios: Orchestrating communities of practice to advance research training. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 220–232.

## Author Profiles

**Daniel Rees Lewis** is a Researcher at Frontiers and the Human-Computer Interaction + Design Center at Northwestern University. He has an MA and PhD from Northwestern University, where he was trained in Delta Lab, the Segal Design Institute, and the School of Education and Social Policy. His research focuses on the design of technologies and services for feedback, coaching, planning for innovation education, including agile, lean UX, and improvement science processes. He led the educational design research in schools, hospitals, and corporations. For 5 years he was the UX Research Lead at ForgeWorks—an educational technology startup he co-founded.

**Matt Easterday** is an associate professor in the School of Education and Social Policy at Northwestern University. He has a PhD in Human-Computer Interaction from Carnegie Mellon University. His research focuses on technology for the new civics – producing scientifically supported educational technology to create informed and engaged citizens who can solve the serious policy problems facing our society such as climate change, inequality, and militarism. He is a Northwestern Center for Civic Engagement Faculty Fellow, a Northwestern Searle Center for Teaching Excellence Fellow, and winner of the Northwestern Ver Steeg Graduate Advising Award. He also serves as Director of Northwestern's Civic Engagement Certificate program, co-founded ForgeWorks—an educational technology startup, and serves as Director of Participatory Budgeting Evanston.

**Chris Riesbeck** is an associate professor of Computer Science at Northwestern University. He is the director of the CS Master's program, and co-director of the Institute for Computer Science and Learning Sciences. He has a MS and PhD in Computer Science from Stanford University. He works in the intersection of computer science, cognitive science, and learning sciences, with a focus on natural language understanding, case-based reasoning, tools for authoring and deploying learning by doing environments, and agile design and development

processes. He is a Fellow of the Association for the Advancement of Artificial Intelligence.

Author Details

**Daniel Rees Lewis**
Frontiers Media | Northwestern University
UK
+44 7481425843
daniel-lewis@northwestern.edu

**Matthew Easterday**
Northwestern University
Annenberg Hall
2210 Campus Drive
Evanston, IL
60208
USA
easterday@northwestern.edu

**Chris Riesbeck**
Northwestern University
2233 Tech Drive
Mudd Room 3117
Evanston, IL 60208
USA
+1 847-491-7279
c-riesbeck@northwestern.edu

Editor Details

**Prof. Dr. Tobias Jenert**
Chair of Higher education and Educational Development
University of Paderborn
Warburger Straße 100
Germany
+49 5251 60-2372
Tobias.Jenert@upb.de

Journal Details

EDeR – Educational Design Research
An International Journal for Design-Based Research in Education
ISSN: 2511-0667
uhh.de/EDeR
#EDeRJournal (our hashtag on social media services)

Published by

**Hamburg Center for University Teaching and Learning (HUL)**
University of Hamburg
Schlüterstraße 51

20146 Hamburg
Germany
+49 40 42838-9640
+49 40 42838-9650 (fax)
EDeR.HUL@uni-hamburg.de
hul.uni-hamburg.de

In collaboration with

**Hamburg University Press**
Verlag der Staats- und Universitätsbibliothek Hamburg – Landesbetrieb
Von-Melle-Park 3
20146 Hamburg
Germany
+49 40 42838 7146
info.hup@sub.uni-hamburg.de
hup.sub.uni-hamburg.de