Adapting Workflow Technology to Design-Based Research:
An Exploration of the Potential of Workflow Expressions to Provide Intermediate
Representations of an Online Learning Environment

Alan J. Hackbarth

Department of Educational Psychology
University of Wisconsin-Madison
February, 2010

A dissertation proposal for the Learning Sciences Area, Department of Educational Psychology,
University of Wisconsin

This Document is Copyright © 2010 by the Wisconsin Center for Educational Research and
Alan J. Hackbarth. WCER retains full copyright ownership, rights and protection in all material
contained in this Document. You may use this Document for your own purposes only. Do not
distribute without permission.
Table of Contents

Table of Contents .............................................................................................................. ii
Table of Figures .................................................................................................................. iii
Abstract ................................................................................................................................ iv

Introduction ........................................................................................................................... 1
Assumptions .......................................................................................................................... 10
Description of the Proposed Study ...................................................................................... 11
Research Questions ............................................................................................................. 12
Context of Proposed Study ................................................................................................. 13
Review of the Literature ..................................................................................................... 14

Intermediate Representations .............................................................................................. 15
Table-based flowchart .......................................................................................................... 15
Descriptive diagrams .......................................................................................................... 17
Conversation maps ................................................................................................................ 18
Discussion ............................................................................................................................. 18

Visual Models of Finer Grain Analyses ................................................................................ 18
CORDTRA ............................................................................................................................ 19
Activity Theory-based Model ............................................................................................... 20
Discussion ............................................................................................................................. 21

Workflow in Business and Science Environments ............................................................... 22
Business-oriented definition ................................................................................................. 23
Workflow Data Mining .......................................................................................................... 27
The Social Infrastructure Framework .................................................................................. 29

Proof of Concept Pilot Study ............................................................................................... 33
Creating the Initial Macro-workflow Expression ................................................................. 34
Creating a Micro-workflow Expression ............................................................................... 38
Redesign of the Macro-workflow Expression ..................................................................... 41

Research Design .................................................................................................................. 46
Context of the Study and Data Sources ............................................................................... 46
Methodological Frameworks ............................................................................................... 48
Design-Based Research ....................................................................................................... 48
Case Study Research .......................................................................................................... 49
Participant-Observer .......................................................................................................... 50

Methodology ....................................................................................................................... 51
Sources of Data Mining ........................................................................................................ 51
Activity logs ......................................................................................................................... 51
Session modules .................................................................................................................. 52
Tool logs ............................................................................................................................... 52
Face-to-face environment ..................................................................................................... 52

Creation and Presentation of Initial Workflow Expressions ............................................... 53
Creating workflow expressions ......................................................................................... 53
Presentation and discussion of workflow expressions ......................................................... 54
Documentation of the Design Process .................................................................................. 55

Future Steps of the Study ................................................................................................... 56
Significance of the Study .................................................................................................... 57
References............................................................................................................................................. 58

Table of Figures

Figure 1. A basic workflow unit. ............................................................................................................. 4
Figure 2. Steps in a designed-based research workflow................................................................. 6
Figure 3. An intermediate workflow representation of the ‘Adventures in Argument’ lesson. ... 9
Figure 4. Sample of a table-based flowchart. .................................................................................... 16
Figure 5. Examples of descriptive diagrams..................................................................................... 17
Figure 6. An example of a CORDTRA diagram. ................................................................................. 20
Figure 7. Wortham’s (2008) coded activity record......................................................................... 21
Figure 8. Graphical comparison of business and scientific-oriented workflow expressions...... 25
Figure 9. Screenshot of the HAL ‘Adventures in Argument’ lesson module................................. 34
Figure 10. A control flow workflow expression of ‘Adventures in Argument’............................. 35
Figure 11. Excerpt of Moodle activity log for ‘Adventures in Argument’ lesson......................... 36
Figure 12. Example of threaded discussion...................................................................................... 40
Figure 13. A micro-workflow expression for the discussion task in the ‘Adventures in Argument’ workflow. ........................................................................................................... 41
Figure 14. Redesigned three-section macro-workflow................................................................. 43
Figure 15. Moodle ‘Reports’ search interface. ................................................................................. 51
Abstract

A fundamental challenge of design-based research is that there are many variables, and interrelations among variables, that affect success of a designed learning intervention. Designers collect large amounts of data, but lack of time and resources often limits analysis and makes conclusions less certain (Brown, 1992; Collins, Joseph, & Bielaczyc, 2004). Researchers in the Learning Sciences have developed various types of intermediate representations to organize data and provide a holistic view of informal and formal learning environments (Angellino, Rogoff, & Chavaj, 2007; Ash, 2007; Barron, 2003), which in turn facilitates more focused, deeper analyses. Workflow technology - utilized in business and applied science environments to describe and automate work processes, reveal “know-how” that is often tacit in scientific processes, and facilitates multiple levels of reuse (Georgakopoulos, Hornick, and Sheth, 1995; De Roure, Goble, and Stevens, 2008) – may provide a useful framework for creating intermediate representations of complex, multi-dimensional learning environments. The core features of workflow technology – i.e., organizing the data, tools, participants, activities, and flow of a system – seem well suited to the core challenges of design experiments – i.e., accounting for a large number and range of variables and their relationships, and the manipulation of variables that occurs during each phase of the design process. Bielaczyc’s (2006) Social Infrastructure Framework provides an initial specification of the critical variables to include in the representation, and principles of information design (Tufte, 1990, 1997) provides guidelines for visually representing the complexity and dimensionality of a learning environment as a web document in two-dimensional space.

I am proposing a study to design and test the efficacy of workflow technology to: 1) address the challenges of accounting for the large number and range of variables and their
intertwined relationships in a design experiment; 2) assist researchers taking a design-based research approach in achieving control in the manipulation of variables, which occurs between implementation phases of the design process; and 3) provide sharable models of interventions that are adaptable to local contexts.

The context for my study is HAL Online, an innovative experimental section of “Human Abilities and Learning,” a required course for teacher education and other education-oriented majors offered largely online through the Educational Psychology Department at the University of Wisconsin-Madison. In online learning environments, many design variable values and their relationships can be recovered post hoc from the data modules and the activity logs of the online course management system. Thus I will also explore potential and limits of data mining for generating workflow expressions that can support design and sharing of online interventions.
Introduction

Design-based research is a relatively new research paradigm that addresses the challenge of systematic design and study of instructional strategies and tools (Barab & Squire, 2004) grounded in the “blooming, buzzing confusion” of classroom learning environments (Brown, 1992) or similar authentic contexts whose complexity can make “systematic study” a daunting prospect. There are many variables that influence the implementation of a design experiment, and many of those variables cannot be controlled (Collins, Joseph, & Bielaczyc, 2004, p. 19) – e.g., the availability of a computer lab at a specific time, or students’ interpretation of a new tool. It is important when doing design-based research (DBR) to identify the critical variables of a design and how they fit together (Collins et al., 2004). In order to evaluate an implementation of an instructional design one needs to analyze each case in terms of its key elements and their interactions. Some elements will be implemented more or less as the designers intended, some will be changed to fit the circumstances, and some will not be implemented at all (p. 34). A profile of how each of the critical elements were implemented and how well the elements worked together toward the designer’s goals is needed. Furthermore, because each variable is part of a systemic whole it is impossible to change one aspect of the system without creating perturbations in others (Brown, 1992). These perturbations, as well as their effect on predicted outcomes, need to be identified and accounted for in the iterative cycles of design, enactment, analysis, and redesign in authentic settings that constitute DBR (Barab & Squire, 2004; Collins, 1992).

It is important to document designs at a level of detail appropriate to the research questions and design goals of the experiment, and critical to record all major design changes at that level, because major changes mark the borders that differentiate one phase of an
implementation from the next (Collins, 1992). The outcome is a detailed design history that characterizes the design elements that are in place in each phase and the reasons for the transitions from one phase to the next. The design history reflects design changes and whether they have any causal impact on measured student outcomes, which lead to theory building and allows research audiences to evaluate the credibility of design decisions and the quality of lessons learned from the research (Barab & Squires, 2004).

However, in the service of documenting and studying design, researchers usually end up collecting large amounts of data during an intervention - e.g., video, student produced artifacts, and activity logs from online interactions - often more data than they have time or resources to analyze (Barron, 2007; Brown, 1992; Collins et al., 2004). To ensure that design researchers maximize their use of the collected data to develop rich representations of critical variables and their interactions in a learning environment, analytical procedures that organize and document data in an easily accessible format that facilitates more efficient deeper analyses are needed (Merriam, 1988, p. 124).

Barron (2007, p. 178) and Derry et al. (in press) discuss the emergence and value of intermediate representations as a response to this need. Focusing their discussion on video research, they argue that intermediate representations are important for identifying what to analyze and for understanding patterns within and across segments of video. Barron identifies several examples of intermediate representations that are either created during data collection or are derived from an initial macro-analysis of video records. These include: content logs, which are often created in the form of field notes that index data while it is being collected; table-based flow charts that chronologically catalog events of a group (i.e., family) informal learning experience and highlight significant events for deeper analysis (for example, see Ash, 2007);
descriptive diagrams of interactions which illustrate participants relations to one another and resources in an informal learning situation (for example, see Angellino, Rogoff, & Chavaj, 2007): and conversation maps that show the flow of discourse in a learning situation and help to identify patterns of differential responding to problem-solving proposals (Barron, 2003).

In these examples intermediate representations provide organizational structures that help the research teams develop a sense of the corpus of data and facilitate the selection of episodes for further detailed analysis (Barron, 2007, p. 179). Further analysis may lead to refined hypotheses (Engle, Greeno, & Conant, 2007), the discovery of new unanticipated phenomena that generates new hypotheses, or design changes that lead to a new implementation cycle.

Macro-level intermediate representations may also be the appropriate level for sharing the details of a successful designed intervention with practitioners. This would satisfy a core characteristic of DBR; that the goal of developing sharable, adaptable learning interventions that are of practical use to practitioners is interwined with the goal of developing theories or “proto-theories” of learning (Collins, 1992, The Design-Based Research Group, 2003). Applied to the structured context of a classroom or online learning environment, standardized models of pedagogical processes may emerge in a manner similar to the way various approaches of “scripting” group interactions (e.g., elaboration, explanation, argumentation, modeling cognition, and question asking) have evolved from the work of researchers in collaborative learning, (King, 2007; Kobbe et al., 2007; Kollar, Fischer, & Hesse, 2006). Standardized models may also help facilitate design by showing at a high level the key elements of a design and their relationships. This would allow researchers to hypothesize about the effect on learning outcomes of changing the order of elements or the content of a particular element and make sure that a change would not drastically alter the intent of an activity or lesson.
An important question to be addressed is, what would the structure of an intermediate representation of a complex, dynamic, multi-dimensional designed learning environment look like? Structure is created in learning environments by setting particular learning objectives, coordinating a series of activities to accomplish those objectives, relating particular resources (e.g., readings, demonstrations, or video) to those activities, organizing workspace, and scripting group interactions to accomplish activities. While a representation of each of these elements would be complex in its own right, a useful intermediate representation would show the flow of activities, the relation of resources to activities, and the way that groups are expected to interact. Such a representation should be feasible to create without an exhaustive expenditure of resources because the basic structure of an educational intervention as a flow of activities makes it amenable to representation as a workflow expression.

The Workflow Management Coalition (WMC) defines workflow in a business environment as “[t]he automation of a business process, in whole or part, during which documents, information, or tasks are passed from one participant to another for action, according to a set of procedural rules” (Hollingsworth, 1995). Broadly stated, a workflow “participant” (i.e., human being or technological tool) receives data in some form as input, acts upon it with a process that may or may not be fully scripted in advance, and its output may become or inform input for the next “participant” (see Figure 1).

![Figure 1. A basic workflow unit.](image-url)
Linking multiple basic units together creates a workflow expression, which may simply describe the flow of information and tasks, or may be used to automate and control all or part of the flow of the system (Georgakopoulos, Hornick, and Sheth, 1995). Workflow expressions vary in terms of detail depending on their purpose; high level conceptual workflow expressions are often used to give general explanations of a system’s function, whereas a significantly more detailed workflow expression is required to analyze processes at an application level (Ludäscher et al., 2005).

Inspired by increasing use of workflow languages to capture, document and control processes in both physical science research and business applications (Ailamaki, Ionnidis, & Livny, 1998; Cardoso, Bostrom, & Sheth, 2004; Ludäscher et al., 2005; Mentzas, Halaris, & Kavadias, 2001), and by trends toward reuse and recombination based on standard workflows and shared databases in the physical sciences (De Roure, Goble, & Stevens, 2008; Gil et al., 2007), I will explore the potential of workflow technology to document a site-based intervention in an online learning environment for the purposes of: 1) supporting design-based research related to incremental development of educational interventions; and 2) sharing adaptable workflows of enacted educational interventions.

My proposed research explores terrain beyond current standard technology tool use in educational research. Most educational researchers are aware that analytical tools such as Transana© and NVivo® provide easily malleable digital workspaces that facilitate organizing, coding and classifying qualitative data, or that statistical packages such as SPSS© and SAS® provide suites of automated tools that allow researchers to organize quantitative data and quickly

---

1 The distinction here between “workflow” and “workflow expression” is between “what happened, or is intended to happen” versus “a diagram of what happened, or is intended to happen”. The terms may be understood interchangeably.
test hypotheses. But literature from business and applied science suggest that an important kind of technology\textsuperscript{2} is embedded within the “workings” of a system and controls the system’s output. Technology-based tools are utilized in business and science applications – and in educational research - to expedite processes that would otherwise be costly in terms of people, resources, and time. More important than the tools, however, is that workflow technology is utilized to organize tasks and tools in a particular sequence, manage the flow of data to and from the tools, and execute the tasks in such a way as to accomplish an overall objective. Often workflow technology is documented as a system is created, \textit{but it can also be discovered, e.g., in the modules and activity logs for a computer-supported system, and exploited to help describe the processes of a system or to learn about how people use a system’s tools, resources, and processes as they interact within it.}

A similar organization of information, tools, people and tasks underlies the processes \textit{and products} of design-based research. A high level workflow for DBR is summarized in Figure 2.

![Figure 2. Steps in a designed-based research workflow.](image)

Importantly for this study, the implementation of a designed intervention in step 2 – a thoughtfully orchestrated collection of learning activities –can be unpacked and represented with workflow expressions created at varied levels of detail. For example, at an intermediate level that

\begin{itemize}
  \item \textbf{Step 1}
    \begin{itemize}
      \item Develop Theoretical Rationale to Guide Design
    \end{itemize}
  \item \textbf{Step 2}
    \begin{itemize}
      \item Implement Designed Intervention in Authentic Context
    \end{itemize}
  \item \textbf{Step 3}
    \begin{itemize}
      \item Systematically Analyze Collected Data
    \end{itemize}
  \item \textbf{Step 4}
    \begin{itemize}
      \item Evolve Design and Theory
    \end{itemize}
\end{itemize}

\textsuperscript{2} In the sense of the American Heritage\textsuperscript{®} Science Dictionary (2002) definition of technology – “the methods, materials, and devices used to solve practical problems” - with an emphasis here on \textit{methods}. 
I’ll call the *macro-workflow*\(^3\) level, workflow expressions facilitate: 1) identification and control of the many variables that intersect during an educational intervention; and 2) sharing of adaptable educational interventions with other educators who can modify the general framework represented by the workflow expression to accommodate local contexts. More detailed views at a *micro-workflow*\(^4\) level (discussed later) provide fine-grained representations derived from data that both constitute and facilitate deeper analysis (step 3) in the iterative process of DBR.

A second related question to be addressed is, *visually*, what would an intermediate representation of a complex, dynamic, multi-dimensional designed learning environment look like? Edward Tufte (1990, p. 33) notes:

“We envision information in order to reason about, communicate, document, and preserve [that] knowledge – activities nearly always carried out on two-dimensional paper and computer screens. Escaping this flatland and enriching the density of data displays are the essential tasks of information design. Such escapes grow more difficult as ties of data to our familiar three-space world weaken (with more abstract measures) and as the number of dimensions increases (with more complex data).”

Visually representing the density, complexity and dimensionality of a learning environment within a limited two dimensional space requires compromise; some elements of a designed intervention need be privileged in its representation, yet techniques must be used that allow access to representations of *all* the critical variables and reveals the logic of their interrelations. Tufte’s writings (1990, 1997) discuss effects and challenges of several such techniques for “enriching the density of data displays”: *micro/macro renderings* in which exquisite detail leads to both focused micro-readings of individual elements of the design and cumulates into larger coherent structures; using *layering and separation* in order to reduce noise and distraction and create relative strata amongst related variables; using *small multiples* for

---

\(^3\) Loosely related to *macro-scripts* (Dillenbourg & Hong, 2008), macro-workflows describe in terms of activities, resources, and tools the environments in which desired interactions are designed to occur.

\(^4\) Like *micro-scripts* (Dillenbourg & Hong, 2008), micro-workflows emphasize individuals’ activities.
comparisons; using color to encode abstract information; and narratives of time and space that account for the temporal and spatial relations among and between elements in a system.

These ideas will inform a rationale that can be consistently applied across different and multiple environments for the visual display of intermediate workflow representations, and that takes into account the affordances of the tools used to create the representations – e.g., HTML documents can be hierarchically “layered” using hyperlinks or encoded using cascading style sheets.

A final related question to be addressed is, what elements (or variables) need to be included in an intermediate workflow representation of a designed educational intervention, and what are their relationships? We know that there are many, and that their relationships are complex and multi-dimensional. We can list some probable ones – students, instructor, small groups, readings, video, worked examples, discuss, view, analyze, etc. Fortunately there are two sources of information in the DBR and workflow literatures that provide starting points from which to explore this question. Katherine Bielaczyc’s (2006) Social Infrastructure Framework (SIF) identifies a number of variables distributed across four dimensions that researchers need to consider as they design learning environments. The Workflow Management Coalition (Hollingsworth, 1995) has identified five key perspectives - what, when, by who, using what data and what tools - that need to be included in a valid workflow expression. These perspectives are defined by many of the same variables found in the SIF.

Figures 3 is an example of an initial attempt at an intermediate workflow representation at the macro- workflow level, generated during an informal “messing about” design pilot study in which I used lesson and activity log data to experiment with ways to visually represent the
activity of a lesson called ”Adventures in Argument” that is part of an undergraduate course in critical thinking at UW-Madison.

Figure 3. An intermediate workflow representation of the ‘Adventures in Argument’ lesson.

This workflow expression was built using the graphical elements of the Information Control Nets (ICN) workflow modeling language (Rembert, 2006), which will be discussed later in the paper. It captures elements of a designed lesson in terms of perspectives such as; what are the tasks being done, in what order are they to be done, by whom, using what tools, and from where is information coming from and going. But the visual display falls short of representing the complexity of a lesson; details such as the goals, pedagogical rationale for the activities, intended interactions of learning materials during activities, or assessments. Some critical variables are not represented, e.g., goals, assessment information, and instructions for activities, and there may be others. Finally, this workflow expression is rendered as a static diagram on paper, which constrains the number and types of visual techniques that can be used to enrich the
representation. I believe that this initial workflow expression can be evolved into a useful intermediate representation of a learning environment by finding the right marriage of workflow principles, information design principles, and critical variables and interrelations that describe the complexity and dimensionality of a learning environment.

I am proposing a study to design and test the efficacy of workflow technology to: 1) address the challenges of accounting for the large number and range of variables and their intertwined relationships in a design experiment; 2) assist researchers taking a design-based research approach in achieving control in the manipulation of variables, which occurs between implementation phases of the design process; and 3) provide sharable models of interventions that are adaptable to local contexts.

**Assumptions**

For the proposed study I make the following assumptions:

- Educational interventions have workflows, which can be expressed visually to serve as intermediate representations of a learning environment that is the product of a designed intervention.

- When an intervention is a product of design-based research, there are, in general, variables or dimensions of the designed intervention for which data needs to be collected for analysis, and these are identified in the design-based research literature.

- For interventions that take place in an online environment facilitated by a course management system (CMS) such as Moodle, many details of the intervention’s workflow are captured in the CMS’s modules and activity logs.

- CMS modules and activity logs can be mined to recover descriptive data about workflow of an intervention that can be represented in a web-based workflow
expression that facilitates; 1) analysis of workflows as part of the iterative process of research-based design, and 2) sharing of adaptable interventions.

**Description of the Proposed Study**

In the proposed study I will:

- Design an approach to create online workflow expressions that incorporate elements/dimensions important to the purposes of analysis and iterative design of a technology-rich educational intervention.

- Develop a limited collection of standard pedagogical processes that are represented in workflow expressions. [This list can be expanded in the future.]

- Develop a method for deriving workflow expressions using as a case a four week unit on children’s thinking in an Educational Psychology Human Abilities and Learning course that incorporates one face-to-face class but is otherwise conducted online (HAL Online). The workflow expressions will be developed from data logged by the CMS (i.e., Moodle), which captures data on both the instructional plans and the students’ interactions, and from data collected in the face-to-face classroom using video research field methods (Derry et al, in press).

- Use the created workflow expressions to facilitate the professor-researcher and her research group in an analysis of the children’s thinking unit through which they will derive design changes to be implemented in the next semester.

- Concurrently solicit feedback from the instructor and members of the research group about the design of the workflow expressions and their utility for analysis, and use the feedback to make design changes in the evolving workflow methodology for design-based research.
• Present online workflow expressions to instructors and researchers in the larger research project in which my study is embedded and solicit feedback on their utility for 1) analysis, and 2) sharing the details of an intervention for potential adaptation and implementation in a similar environment.

• Implement the redesigned workflow methodology in a HAL course that is offered in the same format but with design improvements the following semester.

• Generalize findings to support implementation of this workflow methodology in other face-to-face and online classes that are part of the larger research project in which my work is embedded.

• Draw conclusions about the feasibility of, in the future, automating processes of data mining and the creation of workflow expressions.

Research Questions

The proposed study will be guided by the following questions:

• What elements/variables of a designed intervention are important to represent in a workflow expression of an intervention?

• How is data mined from the CMS-supported intervention converted into a visual workflow expression?

• What level of context of a designed intervention needs to be represented in a workflow expression in order to support analysis and sharing of the intervention? What level of abstraction needs to occur in a workflow expression in order to support generalization (and reuse) of workflow expressions, or elements of workflow expressions? To what extent does a workflow methodology need to balance context of an intervention with abstraction?
Can workflow expressions capture and show differences in interventions that may lead to theory-based explanations of differences in outcomes?

Can the intermediate workflow expressions approach developed in this study be adapted for interventions that take place in other contexts and learning environments, e.g., entirely face-to-face environments or short-duration professional development seminars such as occur in the larger research project?

**Context of Proposed Study**

This proposed study will inform work being done that supports a larger National Science Foundation funded DR-K12 project – Collaborative Research: R&D: Cyber-Enabled Design Research to Enhance Teachers’ Critical Thinking Using a Major Video Collection on Children’s Mathematical Reasoning (hence referred to as the Video Mosaic Collaborative (VMC) project). The VMC project is a collaborative effort between researchers at the University of Wisconsin-Madison, the Robert B. Davis Institute for Learning at Rutgers University (RU), and the Rutgers University Libraries. The project leverages the Video Mosaic (VM), an interactive digital environment supporting video-based teacher professional development and related research. The project is using the VM to support a team of mathematics teacher educators and learning scientists in designing interventions that utilize the Davis Institute’s video collection. This work is being conducted at multiple sites offering in-service and pre-service teacher preparation for different grade levels and mathematics content areas (i.e., counting, fractions, and combinatorics).

The VMC has promised NSF that workflow expressions composed of assessment and intervention data collected independently from sites across the project will be used by researchers and site facilitators to compare and evaluate interventions, supporting the projects’
formative design research cycles. Workflow expressions representing successful interventions are supposed to be incorporated permanently into the VM and indexed to allow future users to locate interventions that are previously tested, have a known probability of success, and are adaptable to local context.

While most sites in the project currently involve face-to-face environments, work has begun on creating online versions of the content area interventions. The extent to which the workflow methodology I will develop in this study is implemented into the design of the VM will be a measure of how effective the resulting workflows are at supporting comparison and sharing of project based interventions.

Review of the Literature

Overview

The concept of an intermediate representation, while not widely touted, is not new in educational research. Initially I will elaborate on three forms of intermediate representations noted in the previous section. I will also review two forms of visual representations derived from tools that support fine-grained analysis of learning environments to illustrate some techniques used by researchers to represent the complexity and dimensionality of learning environments.

The idea of using workflow expressions as a form of intermediate representation is new to educational research. I will use literature from business and applied science to describe the evolution of workflows in both fields and identify the principles that influence their respective designs (Brandic, Pllana, and Benkner, 2006; De Roure et al., 2008; Georgakopoulos et al., 1995; Ludäscher et al., 2005), which will in turn inform the design of educational workflows. I will then review related literature on workflow data mining (Hollingsworth, 1995; Rembert, 2006;
Tick, 2007; van der Aalst, Weijters, & Maruster, 2004), specifically focusing on recommended elements to be included in a workflow expression and processes for recovering and representing data on those elements from application and activity logs.

To close, I will review design-based research literature (Bielaczyc, 2006; Collins, 1992; Collins et al., 2004) for recommendations of critical variables to identify and keep track of during the implementation, analysis, and revision cycles of DBR. In particular, this broad collection of variables is synthesized from a number of design experiments (Bielaczyc, 2001; Bielaczyc & Collins, 1999; Brown, 1992; Collins, 1992) and organized into four dimensions in Bielaczyc’s (2006) Social Infrastructure Framework (SIF).

**Intermediate Representations.**

I begin by reviewing three forms of intermediate representations that have been used to help educational researchers organize collected data into a structure that helps facilitate analysis; **table-based flow charts; descriptive diagrams; and conversation maps.** These methods were used not to analyze designed learning environments, *per se*, but to help each researcher see interesting points or patterns in their data that might promote generalization or be further analyzed.

**Table-based flowchart.** Ash’s (2007) answer to the problem of dealing with the complexity of a learning environment and the consequent overwhelming quantities of data collected from such environments evolved from her research to accurately represent and analyze the actions and meaning making dialogues created by social groups (i.e., families) during their visits to informal science learning settings (i.e., the zoo). To answer the recurrent question of how her research group could generalize from very specific microgenetically-detailed, episode-based analyses, Ash developed a methodology that includes using several levels of analysis in an interlinked way that she describes as moving between levels of macro to micro analyses. The
Challenge her method addresses is to be able to maintain a general overview of events while simultaneously isolating detailed and representative events. The first level, called the Flow Chart (Figure 4), is large-grained and holistic; it provides a chronological overview of one entire visit to the zoo (40-60 minutes) as well as pre and post interviews (15-20 minutes each). The “flow chart” is arranged as a table with headings for time, name of exhibit or event, overview of activity, and specific content or actions during the activity.

![Table-based flowchart example](image)

**Figure 4. Sample of a table-based flowchart.**

Next is an intermediate level called the Significant Event, where one segment of the Flow Chart is analyzed in greater detail, emphasizing dialogue, content, and the kinds of tools the group uses to make sense of the science and connect it to their own prior understanding. The third level is microgenetic and comprises a detailed dialogic analysis of a Significant Event, ideally including details revealed by dialogue, gesture, gaze, and actions. While the Flow Chart is a fairly independent descriptive level of analysis, the identification of the other two levels relies entirely on the research questions chosen and by the theoretical framework in which the researcher works.
**Descriptive diagrams.** Angelillo, Rogoff, and Chavajay (2007) described the evolution of ‘bird’s eye view’ diagrams (see Figure 5) that portrayed the extent and type of mutual engagement in problem solving scenarios between Guatemalan Mayan mothers with varying levels of schooling and three school-age children as they constructed a 3-D totem jigsaw puzzle. Initially diagrams were an informal way that researchers communicated to one another about the roles of participants *vis a vis* group activity. But a number of useful conventions emerged for depicting who was involved with whom and how they were involved, such as acting together, observing, directing others, and playing supporting roles, and the diagrams became useful for analysis and representation of findings.

![Diagrams](image)

**Figure 5. Examples of descriptive diagrams.**

By using one diagram to depict the predominant form of engagement per one minute segment as viewed on videotape, researchers ended up with several pages of small but detailed diagrams for each family’s group session (about 30 minutes). Pages were laid out on a long counter to first simplify the categories and then to examine whether any patterns were evident by mother’s schooling experience (0-2 grades, 6-9 grades, 12+ grades). The clarity of the diagrams allowed researchers to use “eyeball analysis” to understand complex patterns in the data in a matter of only a few hours of examination of the sheets, which were later graphed in more abstracted ways and eventually analyzed statistically.
**Conversation maps.** Once transcripts of an interaction are created, the spatial layout of turns can be designed to make phenomena easier to see. Barron (2007, 2003) describes her own use of a “conversation map” in a study of small group interactions during a problem solving activity. In her study, the turns of each speaker were entered into a unique column and linked by arrows labeled according to an emerging and dynamic coding scheme. The maps hung on her office walls and, while not used to share outcomes of her research, were instrumental in helping her see patterns of differential responding to problem solving proposals that were key to later quantitative coding and qualitative analysis.

**Discussion.** The descriptions of these three studies show how initial representations derived from descriptive accounts or transcription of video data provided an organizational structure of events in each case that facilitated more focused analysis. Collectively they illustrate common characteristics of intermediate representations: 1) a tendency to include visual elements, whether presented as a chronological table, a diagram, or a map; 2) clarity through abstraction – at a core level they capture the essence of the situation, and; 3) the capacity to represent, and allow the researcher to simultaneously manage, varied levels of detail, from the most basic flow of activities to complex interactions of multiple players. As it turns out, these characteristics are also found in workflow expressions.

**Visual Representations of Fine-Grain Analyses**

Music score representations are a common way to show results of analyses when the researcher wants to investigate the interrelations of a large number of variables that describe the environment, or the activity of participants, both individually and as they interact with one another. Here are two examples that illustrate attempts to capture the complexity and multiple-dimensionality of such environments.
CORDTRA. Hmelo-Silver, Chernobilsky, and Jordan (2008) argue that although the primary unit of analysis of Chronologically-Ordered Representation of Discourse and Tool-Related Activity (CORDTRA) is discourse, it can foster holistic visualization of data while enabling fine grain coding and representation of different aspects (i.e., discourse, tools, or structures) of computer-based problem solving environments. In that sense, it would accomplish the same analytical function as Ash’s multi-level model of representation. CORDTRA diagrams contain a timeline associated with lines of a transcript (see Figure 6). Multiple processes are plotted in parallel, allowing a researcher to juxtapose a variety of codes to understand an activity system. (Keyword Maps in Transana™ and the Annotation Board in ANVIL© provide similar functionality.)

However, if a viewer does not know the context of the activity and/or the environment represented by a CORDTRA diagram, understanding the “whole” of the representation might not be possible. Additionally the rationale for selected codes used in the analysis may not be clear from the visual representation. Also, as can be noted in Figure 6, as the number of actors and codes increases, so to does the complexity of the graph; it becomes less clear how an instance of a coded process is related to an actor or how particular instances of some coded processes are related to others.
Activity Theory (AT)-based Model. Wortham’s (2008) model for representing interactions of a small group engaged in a science lab (Figure 7) shares some similarities with Ash’s (2007) Flow Chart; for example, both account for distinct actions across an activity (Ash in a column, Wortham in rows) and both summarize individual contributions and group interactions within cells in the table. Wortham uses a music score model to distinguish individual student’s actions, and devices such as parenthetically reference numbers to represent the flow of discourse. However, the model is laden with the syntax of particular theoretical frameworks, i.e., syntactic elements from Activity Theory (as cited in Engström, 1999), event vocabulary from Goodwin (as cited in Goodwin, 1994), and a specific hierarchical arrangement of levels of action and activity, that constrain the types of questions that can be asked about group interactions.
Discussion. CORDTRA and Wortham’s AT-based model illustrate types of tools that researchers are developing to try and represent the complexity and dimensionality of multi-person learning environments. Rather than producing intermediate representations that are primarily descriptive, these tools produce complete accounts of an event according to the analytical categories chosen by the researcher. The nature of these categories, and the complexity of the representations, is influenced by the tools underlying theoretical framework in one case, and by the research question(s) in the other.
Workflow in Business and Science Environments

The concept of workflow emerged from the notion of process in manufacturing and business systems (Mentzas et al., 2001). According to Georgakopoulos et al., (1995), material processes relate to human-oriented physical tasks (i.e. moving, assembling, or transforming physical objects). Information processes relate to automated tasks (i.e. performed by a computer program) or partially automated tasks (i.e. tasks performed by humans interacting with a computer) that create, process, manage, and provide information. Business processes are typically developed to fulfill a business contract or to satisfy a specific customer need and are implemented as information and/or material processes.

Although contexts and interactions in the world of business are generally more controlled and predictable than in education, a business processes metaphor has some utility in describing the design and implementation of an intervention; instruction can be generally described as a series of processes – including material and information processes – intended to collectively satisfy instructional objectives (i.e. a contract) and satisfy a learners’ (i.e. customers’) needs.

Workflow is a concept closely related to the design, or reengineering, of business and information processes (Mentzas et al., 2001). In a broad and multi-leveled sense a workflow expression is the diagrammatic result of “capturing” the workings of a system from beginning to end. A workflow expression might show any number of possible configurations. For example, a workflow expression may show a business process at a conceptual level necessary for understanding and evaluating a process, or it may represent tasks that comprise information processes at an architectural level that describe the requirements for technology/human interactions (Ludäscher et al., 2005). A workflow expression may be thought of as the set of pre-programmed instructions derived from the study of a system that are automated to run processes.
of the system (Ailamaki et al., 1998), or as a post hoc description of a concrete instantiation of running processes in a particular environment (De Roure et al., 2008).

Within the literature two broad versions of a standard definition of workflow emerge, depending on whether the authors are referring to business or scientific applications.

**Business-oriented definition.**

“A collection of tasks organized to accomplish some business process... A task can be performed by one or more software systems, one or a team of humans, or a combination of these…In addition to a collection of tasks, a workflow defines the order of task invocation or condition(s) under which tasks must be invoked, task synchronization, and information flow (dataflow).” (Georgakopoulos et al., 1995).

This definition elaborates on the WMC (Hollingsworth, 1995) definition presented earlier. *Tasks* hold a prominent position – they are organized and performed to accomplish a process. But specifying the workflow technology of the process – ordering, invocation conditions, synchronization, and information flow - is also a conspicuous part of the business-oriented definition.

**Scientific-oriented definition.** In the scientific community, greater value is placed on access to specialized tools or necessary registers of data that allow scientists to conduct sophisticated analysis or simulations from their desktop computers:

“…[the] user defines the process needed for problem solution as a flow of activities, each capable of solving a part of the problem...Resources that perform these activities are not necessarily located in the user’s vicinity rather they are geographically distributed. This may enable the use of unique resources such as expensive measurement instruments or powerful computer systems.” (Brandic et al., 2006).

A key feature that facilitates the ability of scientists’ to conveniently put together and run their own scientific workflows is the notion of reuse (De Roure et al., 2008). Scientific workflow expressions are not simply digital data objects, *they capture pieces of scientific process* - they are valuable knowledge assets in their own right because they are graphical representations of
“know-how” that is often tacit. This idea of reuse suggests that there may be a standardized way of doing particular processes where the inputs change but the process, or the rationale for the process, remains consistent across instantiations. Reuse can occur effectively at multiple levels: a scientist can reuse a workflow with different parameters and data, fragments and patterns of a workflow can be reused to support science outside their initial application, or they can provide a means of codifying, sharing, and spreading the workflow designer’s practice. Parallels can be drawn here to teachers or designers who introduce new inputs (e.g., data or concepts) while reusing a pedagogical process (e.g., a worked example) from lesson to lesson, and to sharing adaptable interventions.

Comparison of business and scientific workflows. Both business and scientific workflow expressions attempt to define processes of a system in terms of the tasks performed, the order in which tasks are performed, the resources consumed and produced, and the relationships between person and machine. Both account for control flow and data flow. However, when analyzing the underlying principles and execution models of the approaches, Ludäscher et al. (2005) found a focus on control flow patterns and events – what was done, and in what order – in business-oriented systems, while scientific workflow systems tend to have execution models that are more dataflow-oriented – they are interested in how data gets passed through the system (p. 1046).

Visually the difference in business and scientific-oriented workflow expressions is shown in Figure 8. In this case the business oriented workflow expression shows a linear procession of steps and the scientific-oriented workflow shows the hierarchical levels of representation of processes, with each level showing more of the tacit scientific knowledge of the overall process.
Adapting workflow technology to educational settings. When adapting workflow technology to educational settings it is important to think about how to represent control flow (emphasized in business) to model the progression through processes of an intervention (Mentzas et al., 2001), and data flow (emphasized in scientific work) in a way that allows researchers to see how information is accessed and acted upon throughout the intervention (Gil et al., 2007). Educational environments are like both business and scientific environments; at the same time they are unique. As Collins et al. (2004) and Brown (1992) point out, first and foremost is the variability and unpredictability of learners and teachers, who come to an intervention with varied degrees of relevant prior knowledge (Bransford, Brown, and Cocking, 1999). This results in ‘tweaking’ of the processes in an intervention; for example, prerequisite knowledge sometimes needs to be activated (or taught), sometimes additional or alternative conceptual or physical tools need to be introduced to a process, or sometimes groups need to be reconfigured for various
reasons. Veteran teachers often say, anecdotally, that they no longer use a lesson plan because after five minutes they are no longer following it. Instead they follow a general script that is flexible enough to allow adjustments for variability unique to a particular setting. Variability in designs and/or implementation can also be caused by factors that are out of the designers hands – e.g., a specialized room is not available or a person is not available when needed and as a consequence a sequence of lessons must be altered.

This variability and unpredictability suggests that while educational interventions may be initially represented as “intended workflows”\(^5\), they cannot be preprogrammed in the way that many business-oriented workflows are (e.g., loan applications or trip reservations). This further suggests that important functions of an educational workflow include capturing and representing: 1) The intended but adaptable aspects of interventions that can be shared and reused across diverse educational and research settings; 2) The actual workflow of an intervention expressed in terms of the processes students and teacher(s) actually engaged in and if, or how often, they used lesson resources and concepts to engage with one another to accomplish lesson goals, and; 3) detailed data flow within interventions, which supports scientific analysis associated with design-based research. In the first two cases a type of workflow representation that I call a *macro-workflow expression* may be used to represent the designed intervention under study. Macro-workflow expressions can be equated with educational *macro-scripts* – general pedagogical models that aim to create learning situations in which productive interactions and outcomes will hopefully occur (Dillenbourg and Hong, 2008; Dillenbourg and Tchounikine, 2007). A graphical representation of the intended aspects of an intervention form the core of the workflow expression (for example, see Figure 3). Metadata related to the specifics of an intervention may

---

\(^5\) Dillenbourg (European Commission, 2004, p. 3) compares “intended scripts”, i.e., the instructions and/or structures that were designed to achieve a particular kind of interaction, to “actual scripts”, i.e., the task or group interactions that students do actually engage in. “Intended workflows” act in the same way.
be more or less prominent in a workflow depending on if the purpose of the workflow expression is sharing or analysis. If the purpose is analysis, more detailed micro-workflow expressions (equated with “micro-scripts”) may be developed. Much like Ash’s (2007) conception of the “Significant Event” the nature of a micro-workflow is influenced by the research question being asked, the degree of detail a research team requires to answer the question, and the nature of the activity to be analyzed. When studying a micro-workflow expression, researchers might pinpoint a finer grain instance for additional analysis and may choose different analytical techniques to hone their conclusions.

**Workflow Data Mining**

In order to express the variables of a design in a workflow expression, I need a method for mining data on relevant variables from online data modules and activity logs, and for systematically organizing it into a visual model. Much of the literature on workflow modeling focuses on the use of Petri-Nets to automate and analyze business processes (for example, Meena et al., 2005; Tick, 2005; van der Aalst, Desel, & Oberweis, 2000), a highly structured, mathematics-based methodology that describes processes in terms of weighted nodal relationships between places and transitions. The focus on control flow of this method of workflow modeling is too narrow to describe the relevant dimensions of a learning environment. However, the Collaboration Technology Research Group (CTRG) at the University of Colorado – Boulder has developed a workflow modeling language called Information Control Nets (ICN) that is graphical and intuitive and broadens the scope of workflow mining to include a wider range of perspectives than Petri Nets (Rembert, 2006). The primary perspectives included in this modeling methodology include: the **functional** perspective – *what* tasks or activity must occur; the **control flow** perspective – *when* tasks are done; the **informational** perspective – *which* data
is processed and the *data flow* of the process; the *resource* or *organizational* perspective – *who* or *what* performs a task; and the *operational* or *application* perspective – *how* a task gets done (Jablonski and Bussler, 1996; Rembert, 2006; Tick, 2007; van der Aalst and van Hee, 2002). These perspectives are derived from the Workflow Management Coalition’s Basic Process Definition Meta-model – the organization’s definition of what belongs in the specification of a workflow (Hollingsworth, 1995).

In general relevant data is mined from the event logs of the management system that supports the execution of a workflow in order to discover useful knowledge about the processes of the workflow (however in educational environments data can be “mined” from other sources, such as field video and interview data, where logs can be manually created from these sources of data). The ICN workflow modeling language has a mathematical and graphical representation; for the purpose of the proposed study I will focus only on the graphical elements of representation. These elements are shown and described in Table 1:

**Table 1. Graphical elements of the ICN workflow modeling language.**

<table>
<thead>
<tr>
<th>Label</th>
<th>Tasks are represented by labeled circles. A task is an atomic unit of work carried out by one or more people.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Data repositories are represented by labeled squares. Data repositories have production/consumption relationships with tasks, meaning they provide data that is consumed by a task or collect and hold data produced by a task.</td>
</tr>
<tr>
<td>Label</td>
<td>Roles, Which can be comprised of individuals or groups, are represented as labeled rounded rectangles. Roles are often specified in terms of features such as responsibilities, authority, and availability.</td>
</tr>
<tr>
<td>Label</td>
<td>Participants are represented as labeled stick figures. An organizational population specifies which participants belong to which role/group, as well as relationships amongst participants.</td>
</tr>
<tr>
<td>Label</td>
<td>If a task requires a certain <em>application type</em>, e.g., “email”, to be completed, then that task might be accomplished by any <em>application instance</em> of the type specified, e.g., “Gmail” or “Outlook”. An application type is represented with a labeled hexagon; an application instance is represented as a labeled computer monitor.</td>
</tr>
</tbody>
</table>
The are four *control tasks*; two branch tasks - parallel and conditional, and two join tasks – synchronization and merge. *Parallel branches* split between two or more concurrent tasks while *conditional branches* mark a choice between two or more tasks. A *synchronization join* indicates that all tasks directly preceding it must be completed before it can start. A *merge* can begin execution when any of the tasks directly preceding it have completed execution. Parallel branches and synchronization joins are marked with an unlabeled dark circle. Conditional branches and merges are marked with an unlabeled open circle.

The ICN workflow modeling language and methodology developed by the CTRG provides a promising tool for visually representing the activities of a process or system in terms of control flow, data flow, participants and their roles, and tools – in short the kinds of variables identified by Collins (1992) and Collins et al. (2004) as important to account for in design experiments. A more thorough specification of potential variables of interest in educational interventions is provided by Bielaczyc (2006) in her Social Infrastructure Framework.

**The Social Infrastructure Framework**

In introducing the idea of a design experiment, Collins (1992) stated that a long term goal of studying technology innovations in school was to construct a design theory that would; 1) attempt to specify all the variables that affect success or failure of different designs, 2) attempt to specify what values on these variables maximize chances for success, and 3) attempt to identify how different variables interact in creating successful design (p. 19).

The Social Infrastructure Framework (SIF) evolved from research grounded in those initial goals and represents a synthesis of findings from research on learning communities (Bielaczyce & Collins, 1999) and computer supported collaborative learning environments (Bielaczyc, 2001). The SIF is a conceptual tool for examining the social environments and practices, in physical and cyber spaces, associated with designed learning environments.
Adapting Workflow Technology to Design-Based Research (Bielaczyc, 2006). Four dimensions, which are amenable to design and therefore important for comparative study and evolution of learning interventions, are described below.

**The cultural beliefs dimension.** Cultural beliefs refer to the mindset of the teacher(s) and students that shape the way of life in the classroom. Cultural beliefs are not designed per se but they are cultivated over time, and influence such things as: how learning and knowledge are conceptualized, goals (e.g., performance vs. learning), how collaborative/competitive identities of students are shaped, how the identity of the teacher or teachers is understood (in terms of power relations, for example), and how the purpose of technology-based tools is viewed.

**The practices dimension.** This concerns the ways in which teachers and students engage in activities. Of primary interest is data that can reveal how the activities of the intervention operate as a system in which they mutually influence and reinforce one another with respect to overall objectives of the intervention (Brown and Campione, 1996). In this way the SIF aligns particularly well with the idea of workflow, i.e., a focus on the flow of activities in which documents, information, or tasks are passed from one participant to another for action according to a set of procedural rules in order to accomplish a particular objective. Also of interest is how students engage in activities: individually, in groups, or both? Is there a theoretical rationale for how groupings of students are organized or what roles they take on within groups, or how their modes of interaction are supported or constrained (Kollar et al., 2006)? This aspect impacts decisions about the grain size of the analysis of an intervention and how “participants” are represented in a workflow expression. A third area of interest is in the role the teacher plays in the classroom, both in terms of activities and tool use. For example, Bielaczyc and Collins (1999) note a contrast between teachers directing activities versus facilitating student-directed activities (p. 275); the degree to which a teacher balances control with opportunities for
exploration and self-directed learning influences students’ sense of ownership and autonomy. A related critical aspect of the practice dimension concerns whether specifications for activities, participant structures, and tool use are fixed or principle-based, which impacts the level of adaptability students have with respect to, for example, how they work in groups or how they use tools. From a workflow point of view the role of a teacher(s) extends beyond a participant structure. Implementation decisions made by a teacher, often “on-the-fly”, impact what activities or tools get privileged as students “mesh” (Derry et al., 2005) multiple concepts in order to achieve learning goals. A well-conceived workflow expression must be sensitive to the pedagogical moves that a teacher makes in that it clearly shows the degree to which students had opportunities to access materials multiple times, to change the flow of activities, and to re-examine information in light of their emerging understanding of concepts based on their own artifacts

The socio-techno-spatial relations dimension. This dimension refers to the organization of physical space and technology-based workspaces and how they support or constrain teacher and student interactions. Consideration of the physical space needs to be related to how it accommodates anticipated activities and groupings, for example; is there adequate space, can furniture be flexibly rearranged, are mechanisms for sharing information (e.g., whiteboard, projectors) readily available, are there adequate sources of power for technology, how many computers (or other technological tools) are available and how they are accessed, what are the expectations for sharing technology-based information and can that be accommodated by the physical resources? The affordances and constraints of the technology-based tools present another set of issues; how pervasive is the use of technology in the intervention (e.g., very high in an online class and less so in a didactic classroom), how is access provided, how is
information treated in the workspace (e.g., private or public, restricted or open, mutable or static), how are students expected to use tools to interact in the workspace, how is the teacher expected to interact with students through the tools, and how are technology-based activities coordinated with off-line activities?

**The interaction with the “outside world” dimension.** Interaction with the outside world refers to the online and offline ways students are able to interact with people and be influenced by their engagement in events outside their immediate classroom environment. Aspects of interaction to consider include: how is knowledge brought in from “the outside” (e.g., videotape, video-conferencing, the Internet, or invited guest speakers); how is student work extended out to an outside audience (e.g., opportunities for students to actively present or to passively display their work to a community-of-practice or community-at-large in public or online forums); or how are opportunities made available to collaborate with others in/from the outside (e.g., service-based learning projects or “sister schools”)? The door usually swings both ways with respect to the outside world; both students and the “outside” they interact with are influenced by the experience, and that can change the environment on the “inside”. For example, teachers who engage in a professional development course are often asked to implement practices in their classrooms based on the theories they are studying in the course. The interactions often change not only their classroom environments, but also their collegial interactions in the PD course.

**Summary.** With respect to design-based research, the SIF makes explicit many critical variables of classroom social structures that need to be specified and/or monitored when designing and studying technology-rich learning environments. With respect to workflow expressions, the dimensions of the SIF provide a starting point for a specification of elements of
a designed learning intervention. A major challenge in my proposed study is to align the SIF
variables and dimensions with the WMC specification of elements of a workflow (i.e., function,
control flow, data flow, operation, and application) and find techniques to explicitly represent
them visually. This needs to be done in such a way that an instructor can study the workflow and
understand the pedagogy and goals of a lesson, or a researcher can make comparative judgments
(e.g., effect of change in design) regarding two or more implementations of a designed
intervention with respect to the critical variables targeted by the intervention.

**Proof of Concept Pilot Study**

To test the efficacy of the graphical elements of the ICN workflow modeling language to
represent a learning environment, I did an exercise where I recovered archived online data from a
recent offering of HAL Online and created a workflow expression. HAL Online is a hybrid (i.e.,
online plus face-to-face) Educational Psychology course offered at the University of Wisconsin–
Madison that fulfills a program requirement for all education majors and is a recommended
elective for several other programs. It is supported by an information-rich Moodle website that
provides course design tools and online implementation, access to extensive information about
the course (i.e., schedule, instructions, and grading rubrics), access to many course resources
(i.e., videos, web sites, and some readings), and detailed instructions on how to complete
assignments.

I presented the initial workflow expression at a research team meeting and got feedback
on the design which led to two major design moves. The first was to extend the methodology to
include micro-workflow expressions, which would allow researchers to see individual activity
and group interactions within certain activities. The second major change was to reorganize the
macro-workflow expression in order to better represent how resources and activities were
combined into pedagogical structures. I describe the creation of the initial macro-workflow expression, the creation of a micro-workflow model of an online discussion, and the redesign of the macro-workflow in the following sections.

Creating the Initial Macro-workflow Expression

To gather data for the workflow of an online lesson called “Adventures in Argument,” I accessed a module that listed the instructions, resources, and tools for the lesson – including videos to view and a discussion forum. There was also a suggested order in which students could independently view the resources and complete the discussion in small groups (see Figure 9).

---

**Adventures in Argument (Sep 22-28)**

- Pick battles big enough to matter, small enough to win. (Jonathan Kozol)
- Truth springs from argument amongst friends. (David Hume)
- Put the argument into a concrete shape, into an image, some hard phrase, round and solid as a ball, which they can see and handle and carry home with them, and the cause is half won. (Ralph Waldo Emerson)

I. Do by Sep 26 (Thurs):

1. Before reading watch Water Wars videos (links below) and reflect individually on the following: What are the main arguments of the Mesquite NV stakeholders? What are the main arguments of the Beaver Dam AZ stakeholders? What do you think about the quality of these arguments?
2. Read Analyzing Arguments – Halpern chapter 5, pp. 188-228

II. Do between Sep 24 (Thurs) and Sep 28 (Mon):

Required: Participate in your small group forum to analyze Wendell Potter’s argument about health care.

- Water Wars Part 1
- Water Wars Part 2
- Required Small Group Discussion: Analyzing Potter Argument
- Wendell Potter’s Argument
- Bill Moyers Journal: Wendell Potter
- Just for Fun: Argument Clinic

---

Figure 9. Screenshot of the HAL ‘Adventures in Argument’ lesson module posted on a course Moodle site.

Additional data in the form of instructions to students was mined from the online discussion module for the lesson. From these data I was able to construct a control flow...
workflow expression that displayed what was done during the lesson and in what order (Figure 10).

![Diagram of workflow expression](image)

**Figure 10. A control flow workflow expression of ‘Adventures in Argument’**.

Using the element descriptions in Table 1 to interpret this workflow expression we see that at two points in the workflow two tasks were meant to be completed before moving on to other tasks and are preceded by a closed dot for a *parallel branch*. Specifically, two videos are to be viewed before comparing cases, and later two more videos are to be viewed before engaging in a discussion. However in an online environment where students work independently it is difficult to *make* someone watch both videos before they begin reflecting, so the two ‘View Video’ tasks are rejoined by a *merge*, an open dot that indicates the workflow may continue without accomplishing both preceding tasks.

Very quickly I realized that the ICN workflow modeling language did not include elements to represent time (nor did Bielaczyc include it as a dimension of the SIF), or for the type of workspace in which the activities occurred. The latter seems important to represent in hybrid learning environments such as this HAL course that does periodically meet face-to-face throughout the semester. I selected a pentagon-shaped icon to identify workspace and arrowed
lines to represent time (closed-line means ‘must be completed in this time’\(^6\); dashed-line means ‘to be completed in this time frame’).

To determine if this diagram was representative of the actual workflows of the students in the course versus intended workflows (as discussed earlier), I examined the Moodle activity logs for this lesson (see figure 11), which show when individuals open Moodle-based resources, i.e., videos or forums.

![Figure 11. Excerpt of Moodle activity log for ‘Adventures in Argument’ lesson, with names removed.](image)

Using the time that an activity was logged, I could reconstruct an order of activities for each student and determine if the representation of control flow in Figure 10 was consistent with most students. This compliance data is represented as percentages by each activity circle in the final workflow expression (see Figure 3) and show a wide range of engagement by students across tasks. There were two activities for which I could not generate compliance statistics; when students compared cases using the first set of videos (because they weren’t required to upload a record of the comparison to the Moodle site) and whether they read the textbook chapter.

---

\(^6\) However there is no explicit consequence in this case for failing to complete the tasks within the time frame.
(because it is not a resource they would access on the Moodle site). While inferences can be drawn from an extended analysis of the discourse data in the online discussion, my point is that in order to recover direct data about activities or the use of resources, they have to be explicitly embedded in some manner in the CMS.

Using the lesson module and the logs I next added repositories (represented as rectangles) to the workflow expression (Figure 3). Consumption relationships (with arrows pointed toward activity circles) involve watching the videos, which students access as Quick Time files from the Moodle server. A production relationship (with arrow pointed away from the activity circle) occurs during the discussion when students contribute threads and posts to a Moodle discussion forum; these are stored as data on the Moodle server. In the same manner I added elements that represented descriptive data about the ‘Adventures in Arguments’ intervention. For example, the videos that students viewed were served by a digital video application, in this case Quick Time, and the discussion activity took place in “cyberspace” using an online discussion application that is part of the suite of tools available in Moodle. Students watched the videos, compared cases, and read on their own, but participated in the discussion forum in their assigned small groups, with the instructor also participating. Note that I experimented with new icons to represent special kinds of repositories; the silhouette represents prior knowledge – in this case what students know about argumentation – which is activated and “consumed” by the reflection activity. The book icon represents consumable knowledge in the text book.

Dimensions of the SIF are represented more or less implicitly in the workflow expression. For example, from the diagram we might infer that the lesson designer values independent exploration and social negotiation of knowledge. Cultural beliefs will likely always need to be inferred (although surveys or interviews might be used and results made accessible
via some type of ‘Beliefs’ icon), and the workflow representation should make this possible. Individual tasks and the suggested structure of students and instructor engagement with resources and one another is shown, but how those tasks and resources are expected to mesh into a learning activity, how much flexibility students have to navigate between activities, or how the instructor manages a social activity (i.e., discussion) within the learning environment are not at all clear. From the socio-techno-spatial standpoint we know it’s an online environment; we have some temporal information, resources are primarily video served via Quick Time, and the CMS provides an online discussion tool. However, important information about how students gain access to resources and tasks, or use tools should also be available to someone studying the workflow. Finally, we can infer that interactions with the outside world occur via the videos, but information that makes those connections more explicit should also be included.

In summary, the workflow expression in Figure 3 conforms to the WMC specifications for a workflow, but needs several modifications in order to provide sufficient information about critical variables in the designed intervention in order to do appropriate analysis and modification.

**Creating a Micro-workflow Expression**

The discussion with my research group about this initial workflow expression raised two important questions; 1) what level of granularity should an educational workflow expression represent, e.g., implementation level or level of individual students, and 2) how should activities that include social interaction, i.e., discussions or group problem solving, be represented?

A macro-workflow expression, such as depicted in Figure 3, represents the implementation level of a designed intervention, both the intended implementation, which can be discovered in the instructions to students provided in the online modules on the CMS, and
aspects of the actual implementation (i.e., compliance), which can be recovered from activity logs. Macro-workflow expressions facilitate a number of functions: sharing of adaptable interventions; quickly identifying resources, application types and participant structure at a systemic level; quickly identifying compliance of students’ engagement with respect to assigned activities; and control by the designer of the overall implementation when contemplating design changes.

One thing that the team discussed as valuable to design-based research that was not inherent in the macro-workflows, was an ability to see what individual students were doing during the implementation of an intervention, and how students were interacting with one another in collaborative activities. These interactions with course resources and with each other are a usual part of educational designs and may influence modifications that get made to both the intervention’s design and its grounding theories. Data about students’ interactions with Moodle-based resources and activities is available from the CMS activity logs.

To construct a micro-workflow expression that documents the actions of twelve students (organized into two groups) and the professor-researcher during an online discussion, I followed this procedure:

1. Establish a master timeline that spans all of the actions that occurred within the province of the activity.

2. Mine the activity log of individual students to determine which actions each engaged in and in what sequence (*control flow*). Express the sequence of actions on individual timelines.

3. Arrange individual timelines, coordinated with the master timeline, by group as a music-like score.
4. Mine the content of the data logs – in this case the threads of the discussion forum (Figure 12) – to recover data flow; when was new data produced, and from where did discussants get the data they invoked in their postings (or “consumed”)? Data flow is shown with notation (X.Y where X indicates a distinct thread and Y indicates a sequential post to that thread) and dashed lines that connect postings in a thread. [This representation is limited by real estate; for example, the individual timelines don’t show when videos were viewed or the text was read. Connections to readings/videos are broadly represented with tabs, but this richer information would need to be discovered through other kinds of analyses.]

![Image of threaded discussion]

**Figure 12. Example of threaded discussion.**

The resulting micro-workflow expression (Figure 13) provides information about individual students (whose posting activities in an online forum are represented as circles along individual timelines), and a comparison of two groups (separated by a bold timeline). The interactions of two groups are represented in terms of the threads they created during the forums and the posts each student made to each thread. Connections to lesson readings/videos is derived from analysis of discourse and represented with small closed squares. A viewer can immediately
observe stark differences between the groups, interesting temporal patterns in terms of when individual students make posts, and patterns in the way that information in threads is picked up and acted on by different group members. These informal, quick analyses can help direct and focus subsequent deeper analyses, as well as the possible re-design of elements of the lesson.

Redesign of the Macro-workflow Expression

The graphical elements of the ICN workflow modeling language proved useful for creating an intermediate representation of what was done, when, by whom, where information came from and went, and how tasks got done in the “Adventures in Argument” lesson. In other words, it captured the business-like aspects of the learning environment. What the workflow expression fails to capture, however, is the “know-how” that is tacit in the design and execution
of the lesson. Designed lessons have a pedagogical rationale – there are reasons why a teacher has students watch a particular video or study a particular problem solution, or do these activities in a particular order. The value of a workflow expression, in terms of analyzing or sharing the details of a designed intervention, would be increased if it revealed the pedagogical structure of activities in terms of the interaction of critical variables and dimensions of the SIF.

To rethink the structure of an education-specific workflow I went back to the basic workflow unit (Figure 1) – data is input, a process is performed, and data is output. To redesign the workflow expression, I first isolated inputs and combined them according to how they were used in the lesson. Then I identified the processes that were designed to occur in the lesson and described diagrammatically how inputs and activities were combined to accomplish the process. Lessons are often designed so that the outcome of one process (e.g., contrasting cases) serves as input for another process (e.g., online discussion), but it may happen that the understanding of an activity is not completely formed when a student is asked to use it to do another activity. Often students (need to) revisit and redo an activity in order to reinforce or increase their understanding of needed content or procedures. This back-channeling needs to be represented in the workflow expression. Finally summative outputs are produced that may be sources for finer analysis of the outcomes of a lesson, sometimes well after the lesson has concluded (e.g., end-of-unit reflection or quiz). These too need to be explicitly represented.

Taking these potential practices of students and instructors into consideration, I redesigned the macro-workflow expression for the “Adventures in Arguments” lesson. It is shown in Figure 14.
The redesigned workflow expression retains elements of the WMC specification, but I edited several of the ICN workflow modeling language graphics so that they identify specialized types of resources and flow; a video camera icon for video, a book icon for a textbook, a computer screen icon for online instructions, a digital document icon for student-generated documents, a cloud icon for reflective thought, and multi-directional arrows that show the ability...
to go back and forth between activities. These are tentative, of course; I may find as my study evolves that icons need to be more or less specific.

The workflow expression is divided into three sections. The Inputs section shows the specific resources that were used in this designed lesson and where those resources were retrieved from. [Note: a dashed line for an icon indicates that a resource (i.e. video) is housed in an external repository such as YouTube or PBS.] It should be obvious that resources can change from one instantiation of a lesson to the next, in either content or type. In some cases it may be that a designer doesn’t want to, or can’t share the specific resources used in a lesson.

The Processes section is the heart of the workflow expression and depicts the pedagogical processes designed to help students to accomplish the lesson goals. Things that were defined as tasks in my initial workflow (view video, read) are now more accurately depicted as actions that support broader activities (compare cases, online discussion). The pedagogical rationale for the activities is represented more explicitly by the grouping of resources and actions and the use of annotated arrows. For example, in the first of two major lesson activities, individual students watch and mentally compare/contrast two sides of an argument. This activity helps them to focus on the assigned reading and later informs a group analysis of a different argument. But these are not linear activities; students may go back and forth between the assigned reading and either set of videos in order to make sense of or correctly use important concepts. Finally, the instructor is an active participant in the group process of online discussion in that s/he may contribute postings that point to key concepts in the text, the previous activity, or in other group discussions. This relation is also represented by connecting arrows. Note that the Resource icons in the Processes section are deliberately left blank to show that while particular kinds of resources (i.e., video) provide the cases that students compare or analyze, the choice of the actual
videos can be adapted to different contexts. This is a step closer to discovering representations for pedagogical processes that communicate standardized lesson structures regardless of content.

The Output section includes the tangible products of the students’ participation in the lesson. In this example, these include the log of the online discussion, as well as students’ future responses in their reflective blogs and on an end-of-unit quiz. Note that because students were not required to write their comparison of cases in the initial activity, there is no output document. An important function of the Output section is to provide data and the means to do an assessment of the intervention. In Figure 14 sources of data are shown, but there also needs to be a mechanism that allows instructors and researchers to see the actual assessment instruments as they relate to instructional goals, along with scoring guides and rubrics, and actual data about student performance. I envision these mechanisms being provided in an online environment using hyperlinks and pop-ups (activated by mouse-overs) that access relevant hierarchical information about the course, intervention, meeting, and activity.

**Discussion.** These initial workflow expressions represent a “messing about” stage of my work where I familiarize myself with the concepts, tools, and challenges of my proposed study. It is evident that there is a great deal of work to be done – migrating from sketches to design in a web-based environment, creating more comprehensive representations of SIF variables and their interrelations, finding the right balance between context and abstraction, and integrating principles of information design, to name a few. But this work has shown the viability of workflow technology to produce intermediate representations of learning environments, generated excitement in my research team, and helped to identify my next tasks in adapting a workflow expression methodology to DBR.
Research Design

In this proposed study I am exploring the efficacy of workflow technology to inform design-based research of educational interventions, focusing here on a case representing a course taught primarily online with four face-to-face meetings throughout a semester. While there will be useful products that emerge from this research, such as standardized workflow expressions that will facilitate sharing and reuse of interventions in the Video Mosaic repository, the primary questions posed in this study are: Can I invent a standard workflow-based methodology that can be used to detect and document differences in designed interventions, solve problems of control and efficiency in design-based research, and inform theory-based design changes that improve student outcomes in the context of a particular online course that serves as a case study? Can I develop a sound argument that the methods developed for this case study are more widely applicable to the Video Mosaic environment and beyond?

My hypothesis is that such a methodology can improve the effectiveness of design-based research, by contributing a standardized method for generating intermediate representations that reveal useful information about how students and teachers use resources and engage in activities. This method, I believe, will reveal much about practices and/or sequences of activities that are amenable to design changes and might be targeted for deeper analysis, without overwhelming researchers with more data and detail than can possibly be processed within the timeframe of a typical design project, such as the funded project that is supporting this work.

Context of the Study and Data Sources

The context for the proposed study is HAL Online, an innovative experimental Educational Psychology course on critical thinking taught at the University of Wisconsin – Madison. I will be working with a professor and two graduate student researchers who are
conducting a study using a design-based research approach. The study will be conducted in two iterations of the course, Spring 2010 and Fall 2010, and will focus on a four-week unit on children’s thinking. The unit includes four one-week online sessions and one face-to-face class meeting of two and one-half hours. Well developed modules on the course Moodle site provide a list of tasks to be completed for each session and links to most of the resources that students will use when completing the unit. The course typically enrolls 25-30 students who are organized into groups based primarily on their area of study. Students’ grade levels range from sophomore to graduate students. The course is a requirement for education majors and a recommended elective for several other majors; the class regularly includes dietetic, rehab psych, communicative disorders, kinesiology, and psychology in addition to teacher education majors. Diverse geographic and socio-economic backgrounds are represented in the enrollment. I will collect data on two groups of education students, approximately 12-15 students each semester. Random identifiers will be used instead of student names for all data samples produced and distributed to the research team or presented in public.

The usual activities that students engage in during an online session include:

- read a chapter from the course text and/or other assigned readings.
- view a video(s).
- participate in a small group discussion in which they may discuss, compare cases, analyze or argue different positions.
- take a quiz and complete a reflective blog (at the end of a unit only).

In the face-to-face meeting students will work in their small groups and solve a mathematical problem using Unifix cubes. Each group will present and discuss their solutions. As a class they will generalize their outcomes to similar problems.
Data for the study is generated in the four-week unit, but my research work will primarily take place after the unit is completed and will involve: 1) mining Moodle session modules and activity logs and data collected during the face-to-face meeting for a sample of students, and interviewing the course instructor and teaching assistant if necessary, for the purpose of creating initial web-based macro-workflow expressions that will be placed on the VMC project website.; 2) presenting and discussing the workflow expressions with members of my local and broader VMC research teams; 3) potentially generating modified or additional workflow expressions based on feedback/requests from team members to help facilitate further design discussions; and 4) documenting changes in both instrument and course design across meetings and iterations of the unit.

Methodological Frameworks

**Design-Based Research.** A modified design-based research methodology as defined in the Learning Sciences (Brown, 1992; Collins, 1992; The Design-Based Research Collective, 2003) will be employed for this proposed study. I will not be engineering a learning environment, but rather a methodology for deriving a graphical representation of a designed intervention (that informs its evolving design) from data automatically logged by a course management system (i.e., Moodle) as students and the instructor interact with resources and one another in the systems online environment. This methodology is an adaptation of the ICN workflow modeling language (Rembert, 2006) with theoretical grounding in data mining (Fayyad, Piatetsky-Shapiro, & Smyth, 1996; van der Aalst, Weijters, & Maruster, 2004) and workflow management (Hollingsworth, 1995; Jablonski and Bussler, 1996). It is reasonable to assume that as my methodology evolves it will inform the theories in which it is grounded, a core requirement of design-based research, especially given that the methodology is being adapted to
a designed educational environment for the first time. However, the goals of my study are less theoretical than practical; the tool and a related methodology will be my primary contribution.

**Case Study Research.** There are many uses of case studies, e.g., as a teaching device or a form of record keeping, but my interest is as a research methodology that “allows an investigation to retain the holistic and meaningful characteristics of real-life events” while striving to understand complex social phenomena (Yin, 1994, p. 3). Merriam (1988, p. 2) suggests that case study research is an “ideal design for understanding and interpreting observations of educational phenomena.”

Yin contrasts several major research strategies in the social sciences (i.e., experiments, surveys, archival analysis, histories, and case studies) according to three conditions: 1) the type of questions posed; 2) the extent of control an investigator has over actual behavioral events; and 3) the degree of focus on contemporary as opposed to historical events. Case study is an appropriate methodology when “how” and “why” explanatory questions are asked about contemporary events where relevant behaviors of players in the system cannot be manipulated. Merriam (1988, p. 9) and Stake (2003, p. 135) cites a fourth factor in deciding on case study as an appropriate research design – whether a bounded system (i.e., program, person, process, institution, or social group) can be identified as the focus of the investigation. These criteria apply in a related way to the research group’s study of the designed unit of the HAL Online course and to my study of the potential of workflow expressions to serve as intermediate representations for DBR. Stake (2003, pp. 137-138) explains this relations as being an instrumental case study, where a particular case (i.e., a HAL Online class) is examined mainly to provide insight or facilitate our understanding of something else (i.e., workflow expressions as intermediate representations). According to Stake, the case is still “looked at in depth, its
contexts scrutinized, its ordinary activities detailed, but all because this helps the researcher to pursue the external interest.” When an instrumental study extends to several cases, as it will in my two semester study, Stake calls this a collective case study. In this regard, my work will be influenced by case study methodology.

Participant-Observer. Merriam (1988, pp.92-93) notes that there are several stances one can assume while collecting information as an observer, ranging from complete observer to complete participant. In reality a researcher is rarely one or the other. For example, in the proposed study I will act as a participant-observer – “the researcher’s observer activities, which are known to the group, are subordinate to the researcher’s role as a participant in the processes of the group” (p. 92). However, at some point - when the heavy lifting of course design modification is completed - my role must shift to observer-participant, which privileges access to particular information – in this case how other members used workflow expressions to analyze and make changes to the course design.

Participant-observation offers certain opportunities for collecting data (e.g., access to otherwise inaccessible events or groups), the ability to perceive reality from the “inside”, and the ability to manipulate minor events (i.e., schedule a meeting). However, Yin (1994, p. 89) identifies problems associated with participant-observation: the investigator may have to assume positions or advocacy roles contrary to the interests of good scientific practice; the participant-observer may become (too much of) a supporter of the group being studied; or the participant role may consume too much time – the observer-half may not have enough time to take notes or raise important questions. It is important to be aware and account for these potential problems.
Methodology

In this section I will describe three phases of my study: sources of data mining, creation and presentation of initial macro-workflow expressions, and documentation of the design process of workflow expressions as intermediate representations to support DBR.

Sources of Data Mining

I will mine data from the following sources:

Activity logs. The Moodle course management system (CMS) automatically creates a log record every time a person accesses a resource (i.e., video) or a tool (i.e., forum discussion). A ‘Reports’ tool allows a person with administrative rights to access this feature. The interface for the tool (see Figure 15) allows a user to view logs of activity or participation filtered on several dimensions - participant, data, activity, and action. I will “mine” the log for the four week period of the unit for students in the education groups. As necessary I will track when and how often a student accessed each resource (i.e., videos and online readings) and when and how often they contributed to the required online forums.

Figure 15. Moodle ‘Reports’ search interface.
**Session modules.** In addition to the activity logs, the unit or session modules of Moodle may be organized in such a way as to include operational data. That is the case for this course; from the modules I can extract data about the intended flow of activities during the session, the resources that will support those activities, where the resources can be found, and the tools students will use to process or create information. I will start with this information and draw a preliminary sketch of the workflow for a session, and then use data in the activity log to verify the order, or control flow, of the activities. I will also use data from activity logs to verify compliance – how many students actually accessed each resource (see Figure 11). There is at least one activity that I can think of that can’t be verified by the activity logs – reading the textbook. However compliance with this activity can likely be inferred from a more thorough analysis of the content of, for example, students’ forum posts, where they are expected to support their postings with evidence from the course readings.

**Tool logs.** Moodle organizes tools, i.e., forums (discussions), quizzes, resources and wikis, and provides a link to pages in which the data generated by using the tool are listed chronologically in the order they were used/created. For example, clicking on a specific forum link on the Forum page gives access to all of the postings that were made to that discussion forum (see Figure 12). These can be expanded to show initial threads and replies. In addition to forum and quiz data, students regularly publish a reflective blog when they complete each unit. Forum, quiz, and reflective blog data will be mined from these logs when necessitated by inquiries from research team members.

**Face-to-face environment.** Data mining from records and transcripts of the face-to-face environment will contribute to the creation of macro-workflow expressions that describe the activities engaged in by the instructor and students during the meeting and, if of interest to the
research group, micro-workflow expressions of the social interaction activities (i.e., group problem-solving or discussions) that occurred. From the instructor I will obtain copies (or locations online) of the resources used during the class session. I will follow prescribed procedures from “Guidelines for Conducting Video Research in the Learning Sciences” (Derry et al., in press) to videotape the class using two cameras, one from a wide perspective and one focused on one of the education groups as they engage in the problem solving activity. Care will be taken to capture references to physical artifacts (i.e. Unifix cubes) and written representations in the group problem solving activities. The videos will be indexed with field notes synchronized to the video time stamp and transcribed to facilitate analysis. Indexing will be used as a kind of “action log” that will inform the development of an initial macro-workflow expression and transcription will allow researchers to pinpoint sequences of particular interest.

Creation and Presentation of Initial Workflow Expressions

Creating workflow expressions. I will use data mined from online modules and activity logs, and classroom observations to construct online (i.e., web-based) macro-workflow expressions for each session which will be accessible on the local VMC Project website. While there is only one workflow for the unit, I will represent it in five macro-workflow expressions (four online and one face-to-face session) with the understanding that they are chronologically tied together. The design of the workflow expressions will be partially influenced by the initial models presented in this proposal, but more so by a desire to create a design that accounts for the many critical variables and their interrelations as identified in the SIF (Bielaczyc, 2006). My design work will utilize techniques of visual representation (Tufte, 1990, 1997), including micro/macro renderings, layering and separation, use of small multiples, color, and narratives of
time and space to represent the density, complexity and dimensionality of a learning environment.

**Presentation and discussion of workflow expressions.** Members of the research team doing the analysis of the course design include the course instructor and three former TA’s for the course. I will systematically present and discuss macro-workflow expressions of each session with the research group. I anticipate that team members will use the workflow expressions to generate initial observations about the online and face-to-face sessions and as a participant I will offer my own observations. Using workflow expressions and collected data, recommendations will emerge for modifying resources, activities, or pedagogical strategies for the next implementation of the unit. It may also be that standardized representations of pedagogical processes, e.g., worked example, contrasting cases, or jigsaw, might emerge from the group’s discussion and manipulation of workflow expressions.

I also anticipate that there will be questions and suggestions about the workflow expressions and the underlying data, which I will do my best to answer. As an observer I will make note of these questions and suggestions. Data collected as an observer-participant will inform my analysis of the efficacy of the workflow methodology. A challenge of collecting this data is that it will be interspersed with discussion about the efficacy of the unit with respect to learning goals. From a practical standpoint I may find it helpful to audiotape the meeting proceedings and use a transcript to identify statements specific to the utility of the workflow expressions. If necessary, at different points throughout the analysis I will ask for specific feedback on; 1) the efficacy of the workflow expressions to inform the analysis and subsequent redesign of the intervention, 2) recommendations for modifying the designed features of the workflow expressions, and 3) suggestions for modifying procedures for collecting data or
creating the workflow expressions. This information will be integrated into a final design research report that address the research questions.

Additionally, I will actively solicit feedback about the workflow expressions from instructors and researchers in New Jersey. I expect instructors and researchers to be able to review the workflow expression and understand the goals of the lesson, the pedagogy, and how the assessments are related to the activities and goals. Researchers should also be able to examine outputs, see differences between two or more enactments of an intervention, and make judgments or hypotheses about which is better. I will interview these team members or ask them to complete short questionnaires about the intervention to determine what of these details about the intervention they can discover from a study of the intermediate representation. They will also have the opportunity to identify variables that they would want/need to see incorporated into the workflow expressions.

A strong indication of the effectiveness of intermediate workflow representations to facilitate analysis in DBR and sharing of interventions will be the extent to which my methodology and resulting products get incorporated into the VM Intervention tool. The VM is being developed concurrently as I do my work and after two iterations of my study I will know to what degree it will influence the VM Intervention tool (which will be used to share successful interventions).

**Documentation of the Design Process.** Maintaining a design history is an important aspect of DBR as it supports formative reporting on design research. Collins, et al. (2004, pp. 38-39) suggest that there should be five sections to a design research report; consequently an account of the design history should include details collected iteratively of these five areas: goals and elements of the design, settings where implemented, description of each phase, outcomes
found, and lessons learned. I will follow these guidelines in writing my descriptive account of
the evolution of workflow expressions as intermediate representations. Yin (1994, pp. 94-98)
also recognizes documenting design history as part of the data collection and reporting process of
a case study and suggests that a database be established that includes such things as field notes
(and video and transcripts), documents, and tabular materials such as student scores on related
assessments. A procedure is already in place for securely archiving the data generated in a
Moodle-supported course, and for archiving video. A separate database will be created using
NVivo to index and archive all other records associated with the study, i.e., workflow
expressions, transcripts, field notes, etc.

Future Steps of the Study

The primary limitation of the proposed study is its scope – it includes only two design
cycles, it involves only one type of course management system bounded by its inherent data
logging algorithms, and feedback on design changes is obtained from a small number of
educational researchers who have a stake in the designed course. However, this is an exploratory
study that introduces a methodology for data reduction and initial analysis from outside the
educational field and essentially asks the question, “Is this worth exploring more deeply?” As a
part of the research described earlier being done by the Video Mosaic Collaborative, this work
will be expanded in a number of ways. First, the workflow methodology will be applied to create
intermediate representations of the activity in mathematics teacher education courses and
inservice teacher professional development classes and seminars, thereby expanding the scope of
the application. As mentioned earlier, another step would be to work with members of the VMC
to develop generalized representations of pedagogical strategies that can be adapted to particular
contexts, or systematically modified to study effects of various forms of implementation.
Secondly, members of the VM Intervention design team at the Rutgers Libraries will evaluate the design of the workflow expressions in the HAL Online environment for its adaptability to computer-based workflow modeling languages (e.g. BPEL or YAWL). This will impact how it is incorporated into the VM Intervention tool. Finally, VMC members at the Rutgers Libraries will explore development of a system for automatically creating visual workflow expressions from inputted metadata for the purpose of easily sharing adaptable representations of successful interventions.

**Significance of the Study**

As noted in the introduction, a major challenge of design-based research is dealing with huge amounts of data that gets collected across a large number of variables. A procedure is needed for creating intermediate representations for the purpose of helping researchers see significant events and patterns in the complex interactions that take place between people and tools in authentic settings. In the business and applied science world workflow expressions fit such a descriptive purpose (in addition to facilitating the automation of some processes, e.g., bank loans or mineral detection.) By combining data mining techniques and graphical elements of a workflow modeling language (Rembert, 2006) with descriptors of key elements in designed learning environments (Bielaczyc, 2006), and using proven information design techniques to enrich data density, the proposed study has the potential to contribute educational workflow expressions as a form of generalized intermediate representation that supports more efficient analyses of designed learning interventions.
Adapting Workflow Technology to Design-Based Research

References


