Continuous improvement: How systems design can benefit the data-driven design community
Lomas, James and Patel, Nirmal and Forlizzi, Jodi

Suggested citation:
Designing for Continuous Improvement in K12
How systems design can help data-driven design in complex-technical systems
Data-Driven Continuous Improvement in K12
"Data-Driven Decision-Making"

Slavin et al, 2013
1. Gather assessment data
2. Analyze data to identify problems and their causes
3. Select actions to address those problems

Data-Driven Design in K12 classrooms

Product Data UX

Assessment Data

Instruction & Assessment Product UX

Education Companies

Students

MAGICAL IMPROVEMENTS OCCUR HERE
What is Data-Driven Design?

1. Define “the right” goals (or needs).
2. Define “the right” measures of those goals.
3. Take actions to modify the current system when goals aren’t met.
Software Model of Data-Driven Continuous Improvement

Product Users

Software Product System

Product Design Optimization

Product Data Reports

Product Team Reflection
How Might We Design Systems to Support Data-Driven Continuous Improvement?

Simple Heuristic Advice: just make it easier to measure and modify

Measure Outcomes

“What are the areas of need?”

Continuous Improvement Loop

Modify Designs

“What can we do about it?”
Online Product Experiments (A/B Tests): A Method for Continuous Improvement

Which design has the best effect on outcomes?

Over 10,000/day run by Google, Facebook, Amazon, etc.
Continuous Improvement in K12
"Data-Driven Decision-Making"
Slavin et al, 2013
1. Gather assessment data
2. Analyze data to identify problems and their causes
3. Select actions to address those problems

Why is there desire for “Data-Driven Decision Making” (DDDM) in K12?

Very few interventions have been found to be effective in addressing the poverty achievement gap.

In a review of 196 randomised field experiments, Data-Driven Decision Making was found to be one of the most effective approaches for improving high-poverty schools (Fryer, 2017).


DDDM is not “disruptive” or radical change.

DDDM is all about small incremental improvements… ideal for complex social-technical systems
K12 Education is a Complex-Technical System
K12 Education is a Complex-Technical System

- Family Interactions
- Computer Interactions
- Teacher Interactions
- PLCs
- Government Policy
- Out-of-School Programs
- Text Interactions
- Small-Group Interactions
- Administrative Interactions
- Education Companies
The Modern Classroom

- Computer Interactions
- Teacher Interactions
- Text Interactions
- Small-Group Interactions
Small-Group Interactions

Computer Interactions

Teacher Interactions

Small-Group Interactions

Text Interactions

PLCs

“Professional Learning Communities”

• Curriculum Coordinators
• Schools
• Districts
• State Admin

Administrative Interactions
Computer Interactions

Teacher Interactions

PLCs

Government Policy

Text Interactions

Small-Group Interactions

Administrative Interactions

Education Companies

- Learning Standards
- Textbooks
- Assessments
- Paper & Digital Products
K12 Education is a Complex-Technical System

Opportunities from Systems Design?

- Family Interactions
- Computer Interactions
- Teacher Interactions
- PLCs
- Government Policy
- Out-of-School Programs
- Text Interactions
- Small-Group Interactions
- Administrative Interactions
- Education Companies
Systems Design
Role of data in systems design

• Systems use data as feedback to guide action.
• Measurements provide data to inform systems whether goals have been achieved.
• Cybernetic systems?
Mastery Learning
as a cybernetic system

Mastery Learning is a cybernetic system that involves the following steps:

1. **Teach**
2. **Assess Mastery**
   - If mastery is achieved: **Success** → **Next Topic**
   - If mastery is not achieved: **Fail** → **Teach Again**

The system is designed to ensure that the learner masters the material before moving on to the next topic.
Closed System
TEST, OPERATE, TEST, EXIT
Miller et al, 1960

Operate → Test for Goal → Exit
Operate → Test for Goal → Fail
Success

**Open System**

Miller, 1960

- **Controller**
- **Goal**
  - **Operate**
  - **Test for Goal**
    - **Success**
    - **Fail**
- **Exit**
Single and Double Loop Learning

Argyris 1990

Data-Driven Design in K12 classrooms
PLCs

Administrative Interactions

Product Data UX

Assessment Data

Teacher Actions in Response to Data?

Instruction & Assessment Product UX

Product Team Actions in Response to Data?

Education Companies

Admin Actions in Response to Data?

Students

Paper
Case #1:

Supporting Data-Driven Improvement in Educational Product Companies
Common Barriers to DDDM
And some root causes

No Data or Invalid Data
- Product doesn’t measure
- Users don’t implement

Data is Inaccessible
- Poor Data Experience

Data is not Actionable
- ?

No Motivation for Data-Driven Improvement
- ?

Classroom System
Gather Data
Actions for improvement
Analyze Data
"Backwards Design” Methodology
(cite Understanding by Design)

What are the educational goals?

How do we assess we’ve achieved our goals successfully?

What instructional designs might improve the assessments?
Key Question #1:
Is the system capable of measuring success?

Can the system measure whether it is achieving its goals?

“If we want to report on student mastery for topic A, we need to collect valid test data on topic A”
Case #2:

Development plans for a *teacher-facing* recommendation system
Mapping the System to Reinforcement Learning

- What is the space of possible observations? (what we measure?)
- What is the space of possible actions? (what can we do?)
- What is the reward metric? (how do we quantify our goals?)
Mapping Classrooms to Reinforcement Learning

Data from 1000s of other classes

Decisions from 1000s of other teachers

Assessment Data Dashboard

Recommendations

Teacher

Class

Observation

Action Space

Rewards

Decide Next Task

1. Assessment Data Dashboard (Observation)
2. Digital Assignment Possibilities (Action Space)
3. Recommendation Success (Rewards)
Key Question #2:
Is the action space sufficiently defined?

The right data might be gathered and analysed appropriately … but is there a defined action in response?

NEED methods for scaling-up system learning from rich human decision-making
Case #3:

Automated Machine-Learning Optimization of an Educational Game
An AI system for design optimization successfully increased game engagement…

…but then spun out of control!

Multi-Armed Bandit Problem

Goal is to maximize payout from row of slot machines with unknown rates of reward (some machines pay out more)

Need to balance exploration with exploitation

Interface design as a multi-armed bandit problem...

Instructional Design of “Battleship Numberline”

Ship targets
Type to name a fraction

Submarine targets
Click to locate a fraction
Design Factors

**Target Type:** Ship or Sub

**Target Size:** Small to large targets

**Time Limits:** Small to large time limits

**Item Sets:** The fractions to be estimated

**Tickmarks:** Scaffolds for estimating

**Sequencing:** Algorithms for presentation

**Experiment 1:** Assigned 10,832 players to 3 simultaneous experiments, each testing 6 different design factors (2x3: target type and target size)
Success!

Automatic optimization algorithms produced more engagement

Bandit experiments produce greater total engagement
Meta Experiment 1

**Worst Design**
Ship 97

**Best Design**
Ship 90
Meta Experiment 2

Introduced ridiculously large "Bad Designs"
The Data-Experience Dialectic
Data and qualitative insights must be integrated
With 10,000 subjects a day, we could run thousands of experiments each year.

But setting up, analyzing and acting upon experiments is hard.

How might AI get involved? Could we partially automate design optimization and scientific experimentation?
Key Question #3: Is there a human-in-the-loop to keep alignment?

If continuous improvement metrics are not aligned to goals and values, data-driven improvement (teams or AI) will produce unintended consequences

NEED feedback loops about suitability of success metrics
Conclusions
Takeaways

• From Designing Artificial Intelligence to Designing Intelligent Systems
  • Collaboration between teams and algorithms

(not just for AI but for products)
A Definition of Intelligence used in AI

“Intelligence measures an agent’s ability to achieve goals in a wide range of environments.”
—Legg & Hutter, 2007

“an individual’s intelligence is related to their ability to succeed…”
- Robert Sternberg
What can go wrong?

- Poor goals (misalignment with values)
- Poor metrics (misalignment with goals)
- Poor actions responding to data
- Misleading data and misreading data
- Conflicting interests in stakeholders

Unintended consequences!

*Accountability can create perverse incentives*

e.g. schools encourage low-performing students to drop out or to cheat (Schildkamp et al, 2012).
How can system designers help build the human into the loop?

- UI for User Goal Setting
- System mapping & communicating future systems
- Processes to Negotiate Metrics for Success

(not just for AI but for products)
Figure 6.2
Elementary control system
Reinforcement Learning

AI Definition of Intelligence

“Intelligence measures an agent’s ability to achieve goals in a wide range of environments.”
Perception Action

Evaluate Goals vs Metrics

Modify

1st Order: Closed System

Evaluate Values vs Goals

Modify

2nd Order: Learning System

Measure

Goal

Evaluate Goals vs Metrics

Measure

Modify

Action

World

Perception
MGSC 001

Perception

Measure

Modify

Goal

Compare Goals vs Measure

World

Action

Closed System