RESEARCH CASE STUDIES

Brooke Houck, *PhD*

Six studies: four from my work in healthcare licensure and certification, and two from products I built. Each one began with an unclear question and ended in a decision someone had to make.

Exam content, participant identities, and internal data are confidential, so these are summaries.

brookehouck@fastmail.com / linkedin.com/in/brookehouck / portfolio.phronesislabs.net

01

The American Board of Radiology · **Qualitative interviews**

Examiner interviews: standardizing an oral exam platform

THE QUESTION	The board delivers and scores its oral certifying exams through a software platform. Where could a change to the software, or to the work around it, make the exam more standardized from one examiner to the next?
PARTICIPANTS	Medical physicists who serve as volunteer oral examiners.
METHOD	One-on-one, semi-structured interviews. I wrote the questions in advance and left room for whatever else each examiner wanted to raise. The interviews followed the examiner's whole day, not only the screen: • the preparation before using the platform • delivering and scoring exams in it • the wrap-up at the end of each day of testing
WHY THE WHOLE DAY	A tool is used inside a routine. I asked about the work before and after the platform as well as the platform itself, so that a fix could land in the workflow when the software was not the cause.
WHAT CHANGED	*The study led to substantive changes to the software.*

02

Cairn Careers, a Phronesis Labs product · **Product discovery, three audiences**

Cairn Careers: career planning for students, offered through their college

THE QUESTION	Cairn Careers shows college students how exposed a career is to AI, task by task, beside pay and projected growth. Who needs that enough to adopt it, and what do students do with it once they have it?
AUDIENCES	Three groups use the product and each sees something different: students, the career advisors who work with them, and the college that provides it. Discovery so far has been with a university career specialist.

METHOD	A discovery conversation run from a written script, with most of the time kept for the specialist's students and little for the product. • I asked what had happened to specific recent students, not for an opinion of the idea. • I showed the product only on a sample profile, labelled as a sample. • The call was recorded and transcribed, and the follow-up was checked against the transcript.
WHAT I HEARD	Students with real AI project experience list AI as a technical skill on the resume and stop there. They do not say how the project improved a process or created value, and they do well on online assessments without moving on to interviews.
PRIVACY AND ACCESS	The report a college receives works from what the whole cohort does and hides any group smaller than five. Student records are protected by FERPA, so the scoring engine makes no AI calls. I scanned all 23 public pages for accessibility in English, Spanish, and French, and fixed what the scan found.
WHAT CHANGED	*Cairn is now offered to colleges through their career centers, and students join through their own school. The site promises nothing that is not built: no launch date and no price for individuals.*
STATUS	Early stage. There are no pilot results yet.

03 Redesigning a national clinical skills exam

National Board of Examiners in Optometry · Mixed methods

THE QUESTION	What should the national clinical skills exam for optometry (Part III) measure, given how optometrists practice now?
PARTICIPANTS	Practicing optometrists nationally, and committees of clinicians.
METHOD	A sequence in which each stage narrowed the one before it: 1. Town halls with the profession. 2. A national practice survey built on forced-choice pairs. Optometrists chose between exam topics two at a time, which placed every topic on a single scale of how strongly the field favored it. Topics well below the line came out. For topics near the cutoff, I looked at the size of the gap to the next topic before deciding. 3. Clinician content committees, who turned the survey results into exam content. 4. Validity studies during the build, using exploratory and confirmatory factor analysis.
WHY FORCED CHOICE	When a survey asks clinicians to rate topics, nearly every topic comes back "important." Making them choose between two forces a ranking the committees could act on.
WHAT CHANGED	*The work produced a framework of five competency domains and nine clinical presentation categories. The exam grew from four stations to twelve.* *I designed the scoring model, which combines standardized-patient ratings with two independent clinician scores per station into one pass/fail decision.*

04 Early warning for physicians in longitudinal assessment

The American Board of Radiology · Quantitative, behavioral data

THE QUESTION	Physicians maintain certification by answering questions over time in the Online Longitudinal Assessment (OLA). Can the board tell early who is on track to fall short, while there is still time for them to recover?
METHOD	I developed a method that detects early patterns in each physician's repeated results, so staff can contact those at risk.
WHAT CHANGED	*The board gained a basis for reaching out to physicians early.*
PRESENTED	"Giving Notice: Using Cauchy Sequencing to Inform Diplomates of Longitudinal Assessment Progress," with Ding, Laszakovits, Campbell, and Short. ABMS Conference, 2023. Selected as one of the top three of more than 30 posters.

05 An online standard-setting tool for physician experts

The American Board of Radiology · Internal tool, built with IT

THE QUESTION	Physician experts set exam passing standards by judging questions one at a time (the Angoff method). The process relied on manual steps. Could it run in one place?
METHOD	I partnered with IT to design and build an online platform linked to the item bank.
WHAT CHANGED	*The platform cut manual steps from the experts' workflow and automated the cut-score calculations.*

06 Business Time Back: honest answers without monitoring

Business Time Back, a Phronesis Labs product · Product and survey design

THE QUESTION	Small business owners and team leaders want to know where the week goes. Time trackers answer that by monitoring people. Can a four-minute self-report give a useful answer instead?
DESIGN	• Each person enters rough weekly hours for their own work, with no timer and no account. • A team leader sends a link and never fills it in for the team. The leader sees the team total by task only once three people have answered, and never who said what. • Nothing is saved until the person says so. • The results are labelled as directional. The tool ranks opportunities and does not claim precision to the minute.
WHAT IT PRODUCES	*A ranked map of the hours that could come back, and one change to make first, with a review date two weeks out.*