Are 109 Participants Really Enough?

Jennifer Lee Carlson  
User Experience Manager  
Cisco Systems, Inc.  
San Jose, CA 95134, USA  
jennicar@cisco.com

Abstract  
One of the three-horned problems that all researchers in the CHI communities from both ACM and AIS must tackle is to optimize the generalizability of findings to a broader population. UX Researchers in Worldwide Sales Strategy and Operations at Cisco Systems conducted a large-scale study seeking confidence and insights to guide the business in investing millions of dollars for capability improvements. The study aimed to learn:

- How well a set of proposed capabilities would solve target sales users’ challenges
- How useful the capabilities would be for the different sales roles
- How likely these users would be to take advantage of the capabilities

We completed a concept validation study by conducting 109 one-on-one sessions with participants world-wide. The results guided the company’s sales digitalization priorities for the next fiscal year.

This paper describes how we designed the research, particularly the participant number target and criteria, in order to generalize the findings to a broader user population, and discusses the challenges we encountered in the process of generalization.

Author Keywords  
User research; UX research; user-centered design; sales enablement; business strategy.
ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction
How can a business team gain enough confidence to invest millions of dollars to improve over ten thousand sales people’s efficiency? Business owners need to understand how well the capabilities they propose would solve target users’ challenges and how likely users would be to adopt these capabilities. The user experience research team in the Worldwide Sales Strategy and Planning organization at Cisco Systems was engaged to gather user insights to help gain such confidence.

We designed a concept validation study targeting 122 participants. Business owners wanted to have a large sample size in order to:
- Achieve sufficient confidence in the data to make critical business decisions
- Understand the different perspectives of the various sales roles, market segments, and regions
- Gain insights to guide the direction of the solutions

Management wanted to see the research results quickly, before their project funding decision deadline. The first research session started on June 1st, 2017 and a quick-hits result was needed by June 19th. Four researchers conducted 109 half-hour one-on-one sessions during a two-week period, using a carefully designed data-logging template. We automated data analysis as much as possible.

Results from the in-depth data analysis provided the level of confidence that the business team needed for overall capability prioritization. However, we did not have sufficient data to generalize trends by any other parameters such as participants’ job roles, regions or target market segments.

The Project: What We Did
The criticality, scope and complexity of the project demanded extra discipline and structure, as well as creativity to enhance our agility. The project included the following steps:
1. Agree on the scope
2. Define participant target and criteria, recruit participants
3. Create role-based research protocols
4. Design a data-logging template
5. Conduct participant sessions
6. Automate data analysis and report the results

Each step in the process had its own objectives and challenges, described next.

1. Agreed on the scope
What level of detail could we consider as “concept”, versus “requirements” and “design”? Initially, business owners wanted to validate many details within a 30-minute session. We used two simple questions to agree on the level of detail:

a) Will this be enough detail for participants to understand the capability concept and provide feedback on the concept?

b) Will it be a waste of time to validate at this level of detail if the capability can’t even solve users’ challenges?

2. Defined participant target and criteria, recruited participants
The business team wanted 122 participants to cover four categories of sales roles across four major market segments, distributed among all regions: Americas, EMEAR (Europe, Middle East, Africa and Russia), APJGC (Asia Pacific, Japan and Greater China). These four role categories and target participant counts were:

a) **40 account managers** (we got 32)
   - 8 from each of 4 market segments: commercial, enterprise, service providers, and public sector
   - 4 from each of 2 acquisition teams

b) **32 technical sellers** (we got 30)
   - 8 systems engineers dedicated to very few large customer accounts and 8 who support a large number of smaller accounts
   - 8 sales specialists dedicated to very few large customer accounts and 8 who support a large number of smaller accounts

c) **24 virtual sales team members** (we got 18)
   - 8 virtual account managers, 8 customer success managers, 8 virtual systems engineers, and 8 virtual sales specialists

d) **26 sales leaders** (we got 29)
   - 20 managers: 6 of account managers, 6 of technical sellers, and 6 of virtual sellers
   - 6 directors, who typically have multiple sales roles within their organizations
   - The extra 3 participants were leaders from the acquisition teams

The participant distribution was based on three factors:

a) **Distribution of the actual user population** among the four role categories above.

b) **Impact of the capabilities** that were evaluated.

c) **Usage of the tool** in which the capabilities would be built – salesforce.com, of which account managers are the heaviest users.

Three field engagement managers from the Americas, EMEAR and APJGC regions recruited a total of 163 participants locally. Regional distribution of the target and final participants matched the actual user population.

3. **Created role-based research protocols**

We decided to conduct one-on-one interviews instead of a few role-based surveys because, in addition to predicting the likelihood that users would take advantage of the capabilities, we also needed to understand why a capability would or would not improve users’ experiences. We developed 3 variations of protocols and scripts:

a) **A general version** to cover all the capabilities except those that were specific to sales leaders

b) **A sales leader version** to cover the capabilities that are specific to them, and to capture their perspective on how well the other capabilities would benefit their team members

c) A version for sellers from recently acquired companies, who may not have experienced some of the legacy challenges

4. **Designed a data-logging template**

To manage inter-rater reliability, we created a data-logging template in Excel, and agreed on how to use it.
5. Conducted participant sessions
Four researchers worked with the 4 categories of roles individually. They completed 109 one-on-one sessions via Cisco WebEx, a video conferencing tool.

6. Automated data analysis and reported the results
The quick-hit results consisted of 66 charts, showing overall average and averages by role categories. They were mostly generated by Excel formulas:

- Percentage of participants who had experienced the challenges as stated
- Percentage of participants who agreed that a capability would lessen the stated challenge
- Level of agreement with statements that the participant would use a capability

Based on the numeric results, the business team decided to focus on 3 of the 5 capabilities.

Discussion
Designing and conducting this user research project was very complex, given the criticality, research objectives, participant target, session time, and schedule constraints. It was important to maintain consistency among the different researchers who worked with users in different categories of roles.

There were 70 combinations of key participant characteristics, including job roles, management levels, target market segments, geographic regions, and legacy vs. acquisition sellers, all of which could potentially contribute to higher or lower ratings of a capability. Even with 109 participants, we could not confidently generalize any possible relationships between so many combinations of characteristics. A higher participant count would increase the total time we took sellers away from selling, increase the cost of conducting the study, and delay the investment decision. Without exception, we as ACM SIGCHI research practitioners in the high-tech industry are challenged when trying to optimize the generalizability of findings to a broader population.

Conclusion
The overall result from 109 data points was sufficient for business owners to confidently prioritize the capabilities. Without this study, there would have been many lengthy debates among the stakeholders, or a different prioritization, which could have led to less desirable outcomes for the target users. This sample size achieved the same margin of error (+/- 9 at a 95% confidence level) as the targeted sample size of 122. However, there were not sufficient data to generalize about factors that might impact prioritization for specific user groups.

Acknowledgements
The author wishes to thank the following managers and team members who contributed to this UX research.

- Eileen Hsu, Director, Worldwide Sales Strategy & Operations, for her sponsorship of this research.
- Daniel Fleischer, Jyoti Bhattacharya and Brendon McLoskey, Business Owners, for emphasizing the importance of a concept validation and sample size, and for making the capability concepts “research ready”.
- Stacey Heimkes, Peter Salisko, and Brian Schindwein, Field Engagement Managers, for recruiting the 163 global participants.
- Lori Eronimous, Elizabeth Hanna, Michael Siepmann, Christina Tsai and Carol Swanson, UX Research Consultants, for conducting and supporting this research project.
References