

The Impact of Advanced Computing Training on Researchers and their Work

Conducted by
The Evidence Network Inc. (TEN)
October 2021



Introduction

The Ontario Consortia

In Ontario, the largest providers of Advanced Research Computing (ARC) resources, including training, are four academic consortia based at universities and hospitals.

These consortia have strategically built and supported ARC platforms since the late 1990s, including two of the country's current five national ARC platforms: Niagara (at the University of Toronto) and Graham (at the University of Waterloo).

The academic consortia play an essential role in providing access to ARC systems, programming expertise, training and other supports to researchers in Ontario and across Canada.

- SciNet is led by University of Toronto and hosts the Niagara system
- CAC, the Centre for Advanced Computing, is based at Queen's University
- SHARCNET is a consortium of 19 universities and colleges and is responsible for running the Graham system, located at the University of Waterloo
- HPC4Health is a partnership between SickKids and the Princess Margaret Cancer Centre at University Health Network

ARC Training in Ontario

- Researchers who wish to use advanced computing in their work need to develop specialized skills that are not typically taught in university or college programs
- *And* researchers need to keep these skills current, as systems, software, research problems and techniques continuously and rapidly evolve
- These critical needs are fulfilled by comprehensive ARC training provided by the consortia, which includes:
 - User training for people with a Compute Canada account. Examples: New user training, 'pre- and post-production on Graham', 'Quick tips for getting the most out of SHARCNET'
 - Introductory or overview courses. Examples: 'Introduction to parallel computing', 'Introduction to machine learning'
 - Courses in programming skills and languages. Examples: data visualization, Python, C++, Modern Fortran
 - Focused exploration of specific topics of interest. Examples: 'Is my neural network too big to fit into a GPU', 'Options for solving jobs with many tasks'

ARC Training in Ontario (cont'd)

- Courses are typically offered at levels from introductory to advanced
- Course format runs the gamut from lunch-time webinars to week-long 'summer school' programs. Many sessions are recorded and made available on YouTube
- The Ontario consortia deliver 27,000 person-hours of training to 10,000 attendees annually, which represents 60% of all the training delivered across the country by the Compute Canada Federation
- Compute Ontario is creating a Training Advisory Committee with representatives from the consortia and other stakeholders
 - Initial focus will be to coordinate the training efforts currently underway across the province
 - Longer term focus will be creating a strategic plan for training in Ontario
- To inform the work of the Training Advisory Committee, Compute Ontario commissioned The Evidence Network (TEN) to conduct a baseline study of the perceived impact of training provided by the four consortia; the results are presented in this report.

Survey Information

- Registration for the survey opened on 14 September 2021. A total of 156 individuals registered
- The impact survey launched with a private link (contact information was required) on 14 September 2021, and researchers were able to respond until 6 October 2021
- To encourage participation, a public link (no contact information required) was provided to registrants and communicated via Twitter beginning 1 October 2021, and researchers were able to respond until 6 October 2021
- A total of 266 researchers responded to the web-based survey
 - 93 researchers responded to the private link survey
 - 173 researchers responded to the public link survey

Key Findings

Key Assessment Findings

- ARC training has significant benefits for researchers, including:
 - 89% reported tangible benefit to research capabilities (i.e., exposure to concepts or techniques, ability to tackle new research problems, or research productivity)
 - 67% reported positive career impact
- Level of training, and training frequency (i.e., # of times) are positively associated with impacts on research capabilities and career progression
- In general, training participants are very satisfied with the training provided and have recommended the training to others in their network

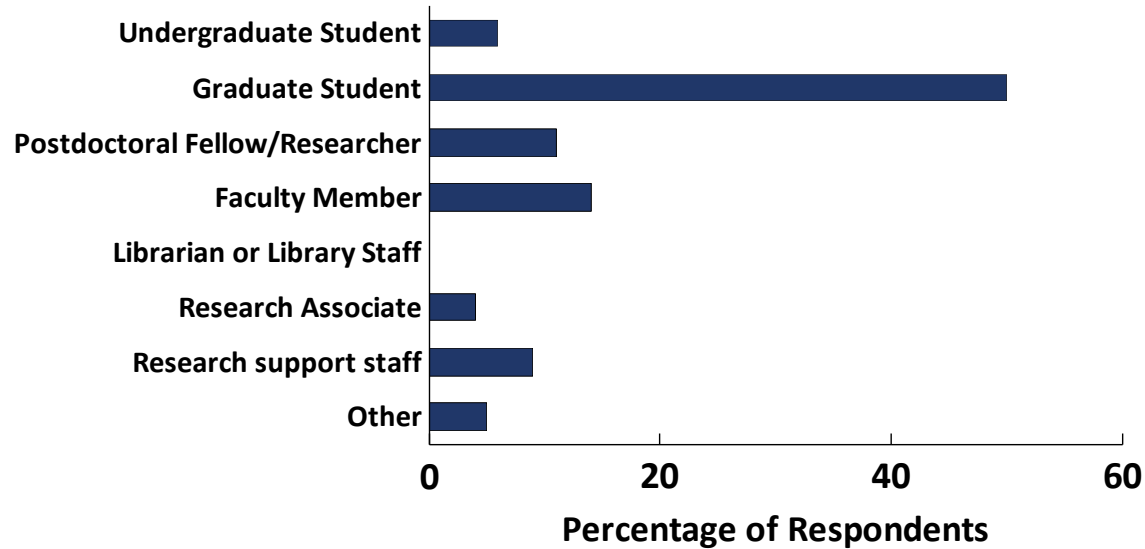
Key Assessment Findings

- 25% of respondents reported that they were affiliated with a university, college, research institute, or other organization outside Ontario, indicating that those outside of Ontario derive value from the training. Of these respondents, 86% received the training in an online format
- More than half of the training participants have received different types of training in a multimodal way
- Training participants are frequently graduate students, affiliated with organizations in Ontario, and with a research focus on the natural sciences

Respondent Demographics

Demographics

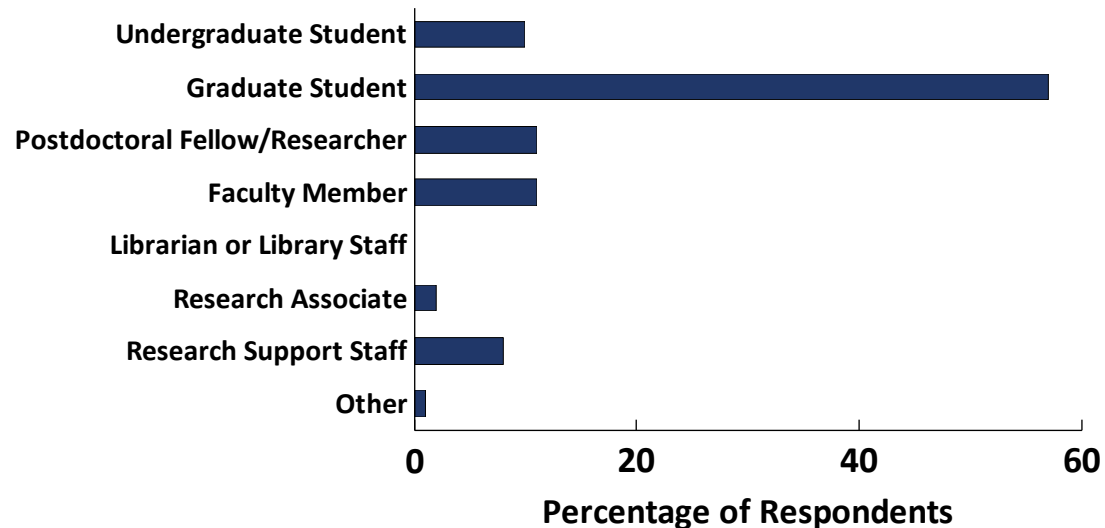
Current Position



50% of respondents reported that they are currently a graduate student.

n=265

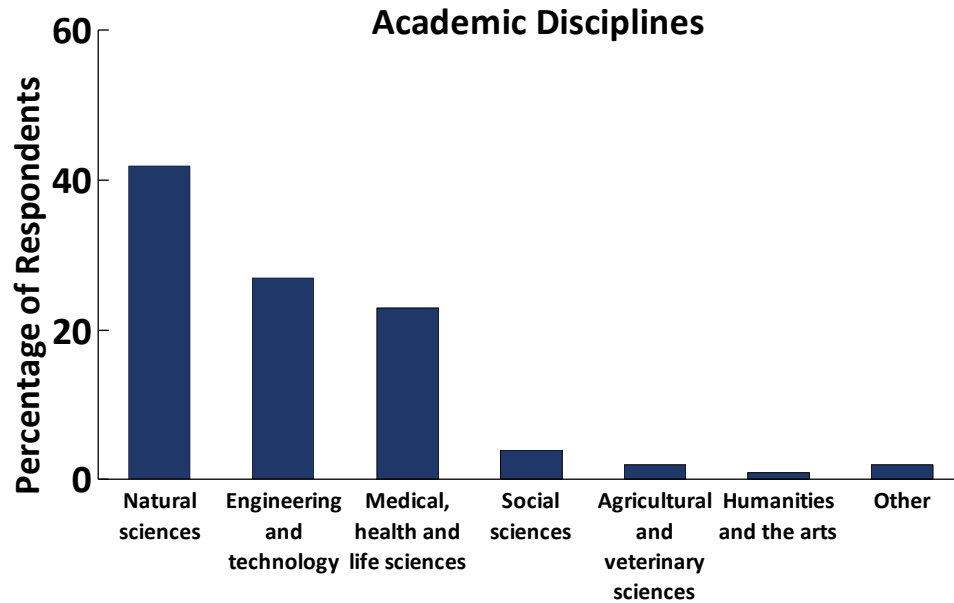
Position When First Received Training



57% of respondents reported that they were graduate students when first received training from the Ontario ARC consortia.

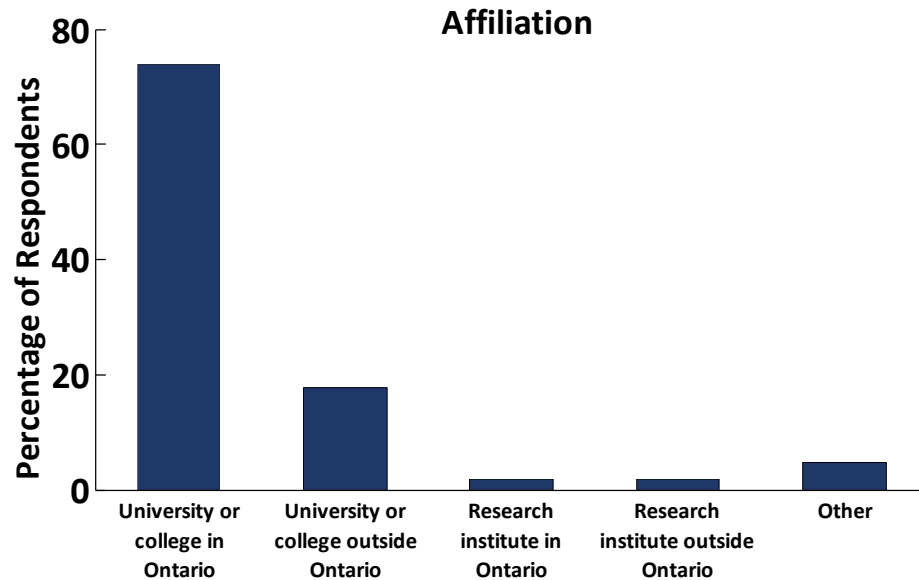
n=264

Demographics



42% of respondents reported that their research focus falls within the natural sciences discipline.

n=265



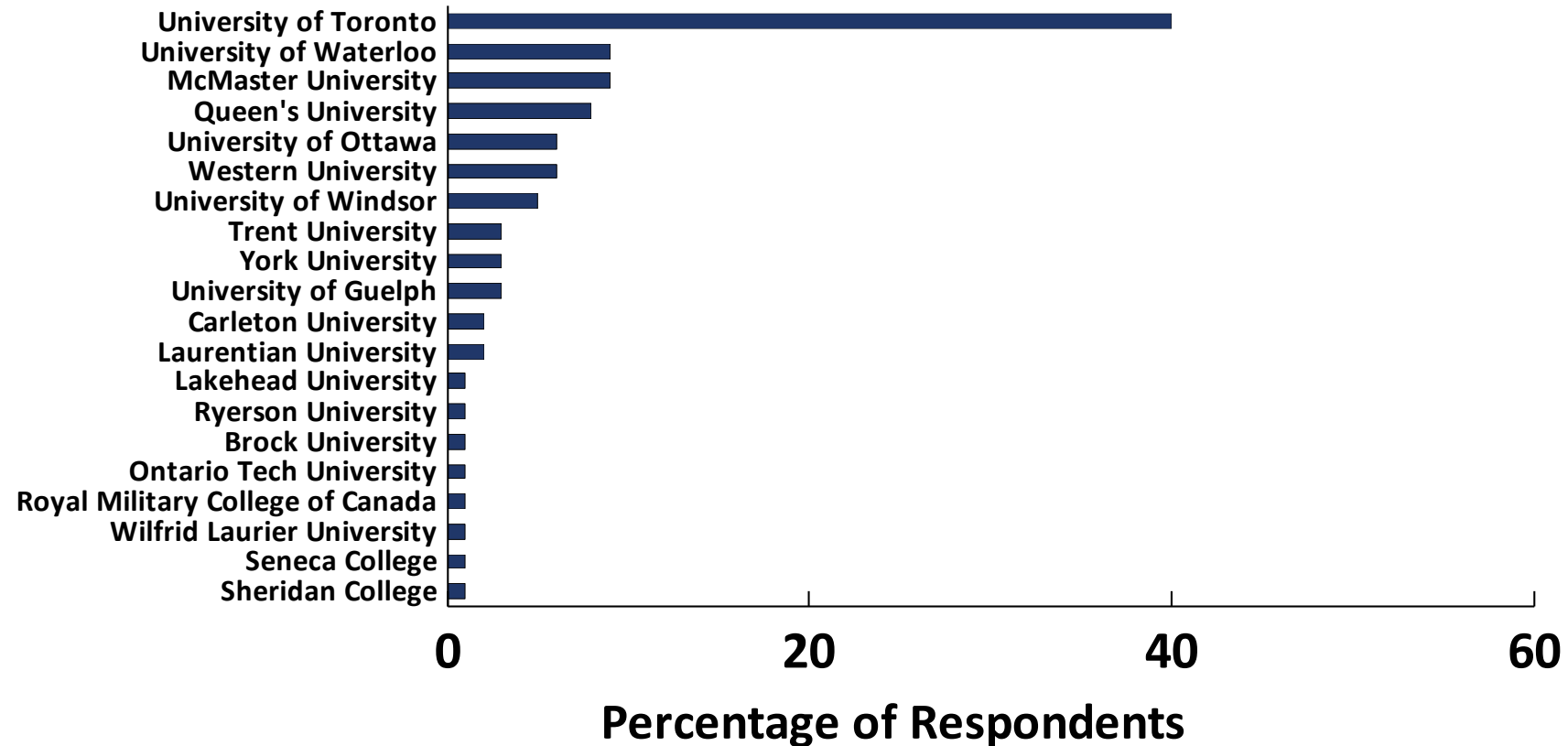
75% of respondents reported that they were affiliated with a university, college, or research institute in Ontario.

25% of respondents reported that they were affiliated with a university, college, research institute, or other organization outside Ontario.

n=265

Demographics

Affiliated University or College in Ontario

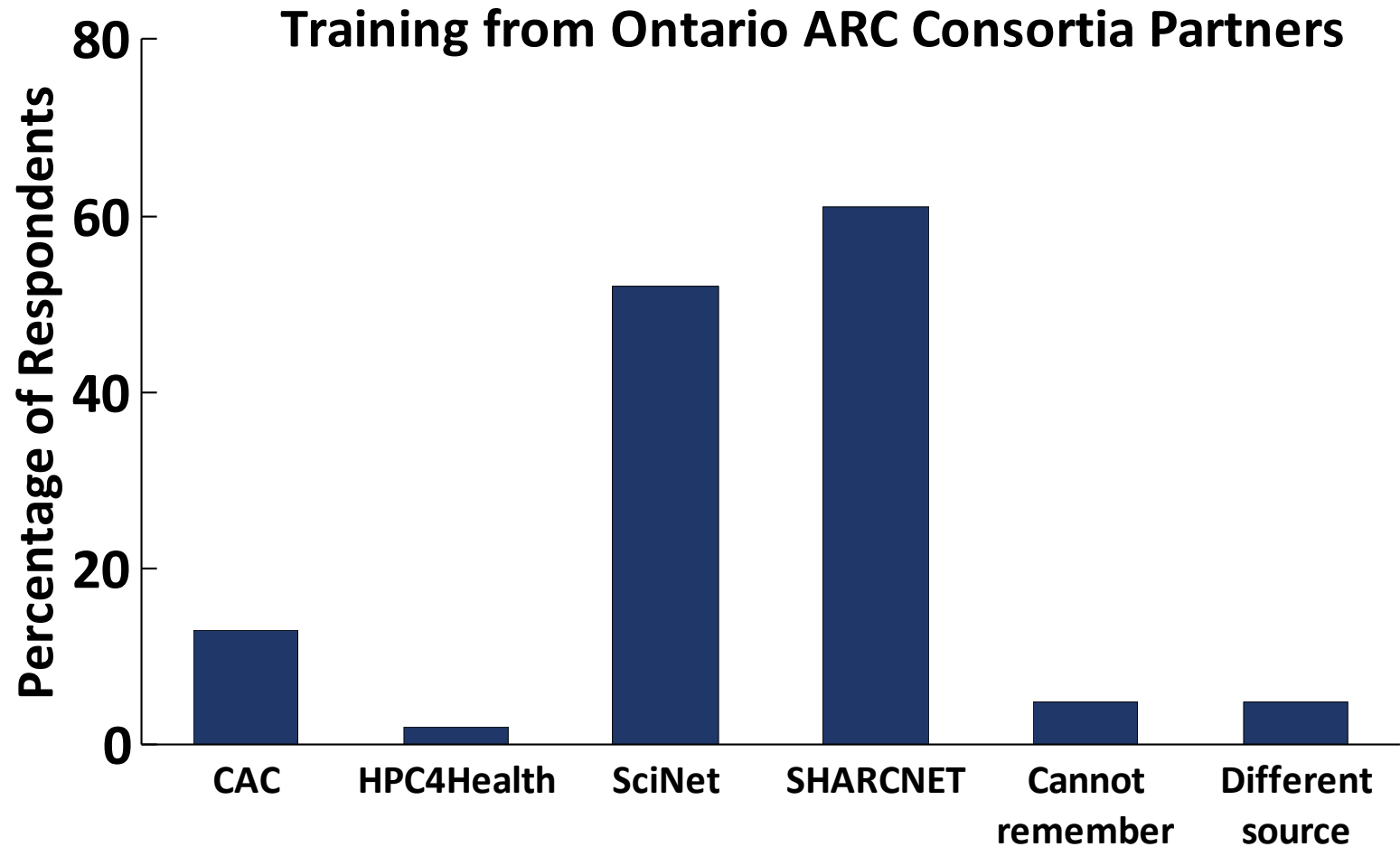


40% of respondents reported that they are affiliated with the University of Toronto, of which 90% reported they have received training from SciNet. Note: 17% of Ontario university students are from University of Toronto

n=196

Training Frequency

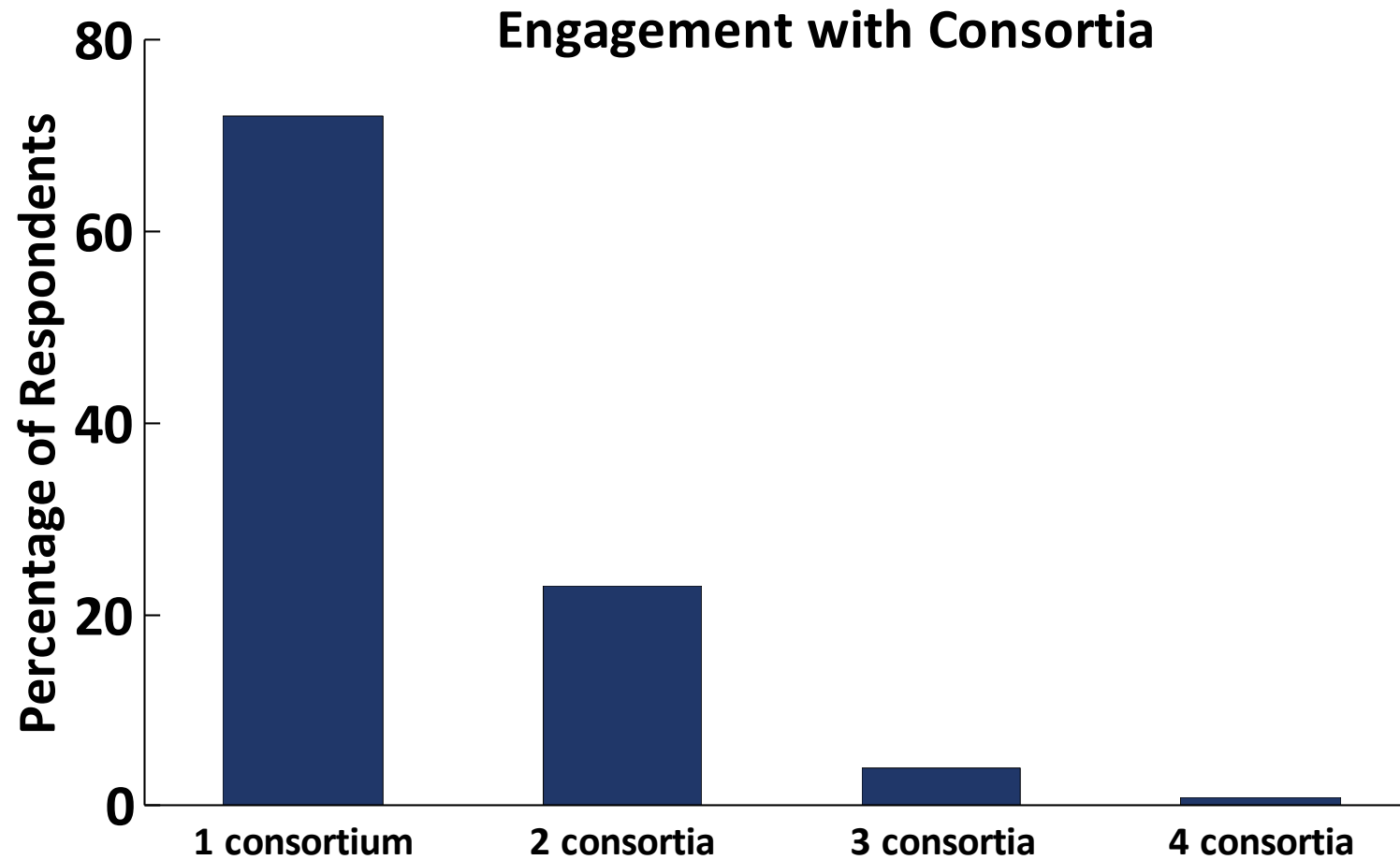
Training Received from Consortia Partners



**61% of respondents reported that they have received training from SHARCNET.
52% of respondents reported that they have received training from SciNet.**

n=256

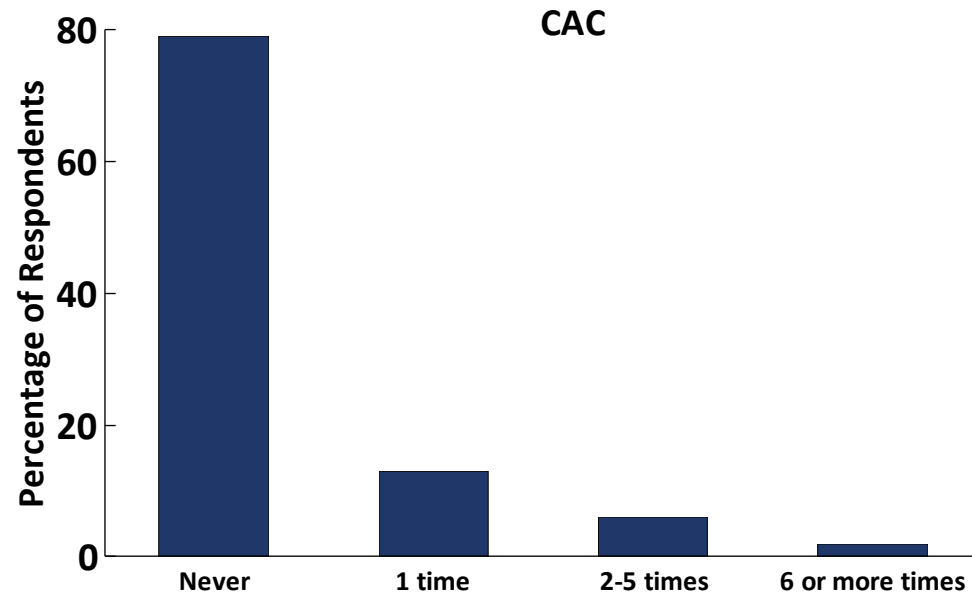
Training Received from Consortia Partners



28% of respondents reported that they have received training from at least two consortia.

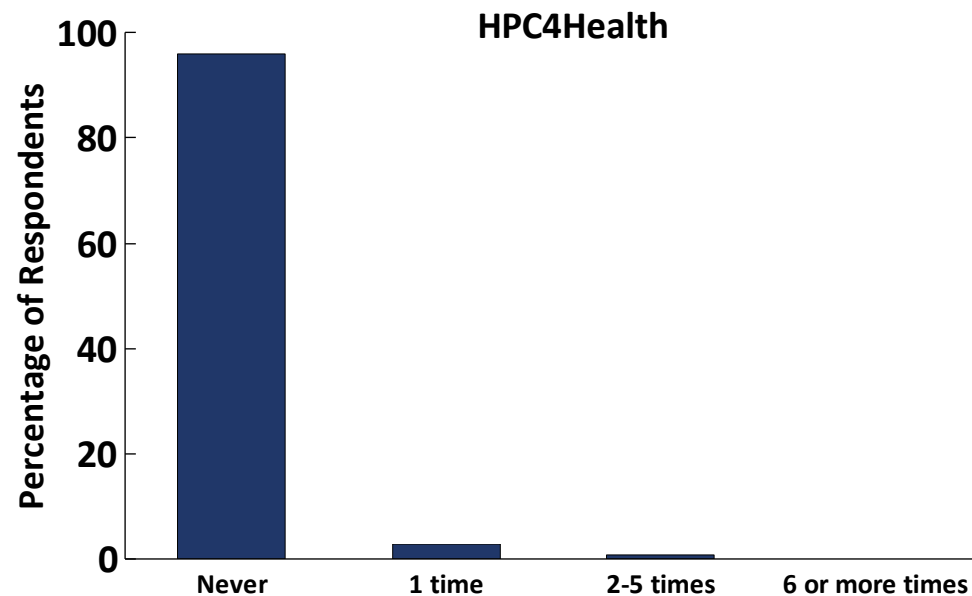
n=255

Training Frequency



21% of respondents reported that they have received training from CAC.

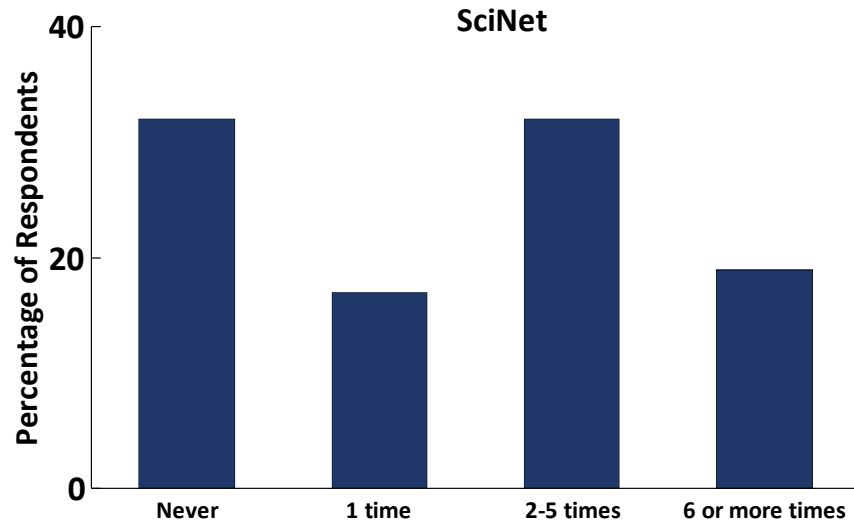
n=154



4% of respondents reported that they have received training from HPC4Health.

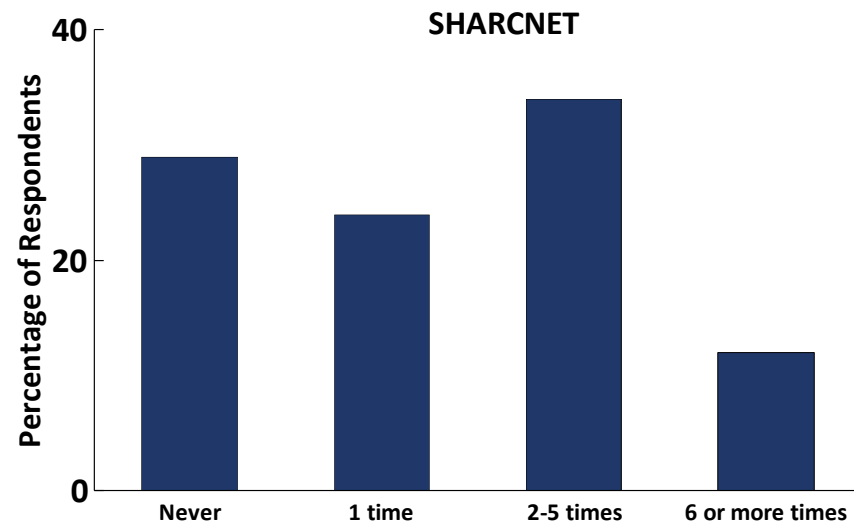
n=144

Training Frequency



51% of respondents reported that they have received training from SciNet more than once.

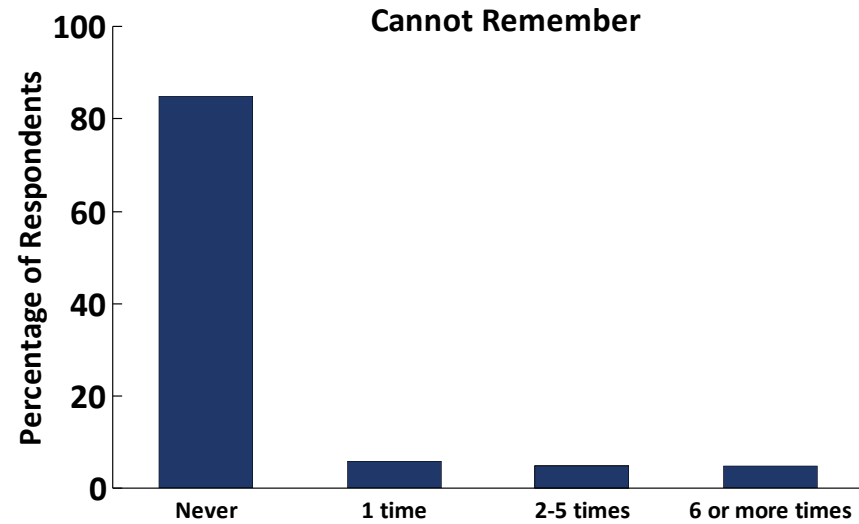
n=196



46% of respondents reported that they have received training from SHARCNET more than once.

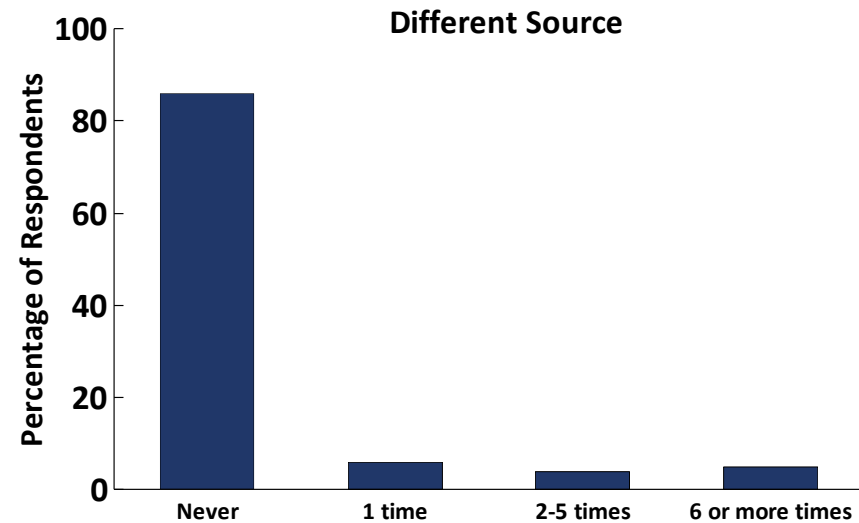
n=221

Training Frequency



15% of respondents reported that they have received training but cannot remember the source.

n=88

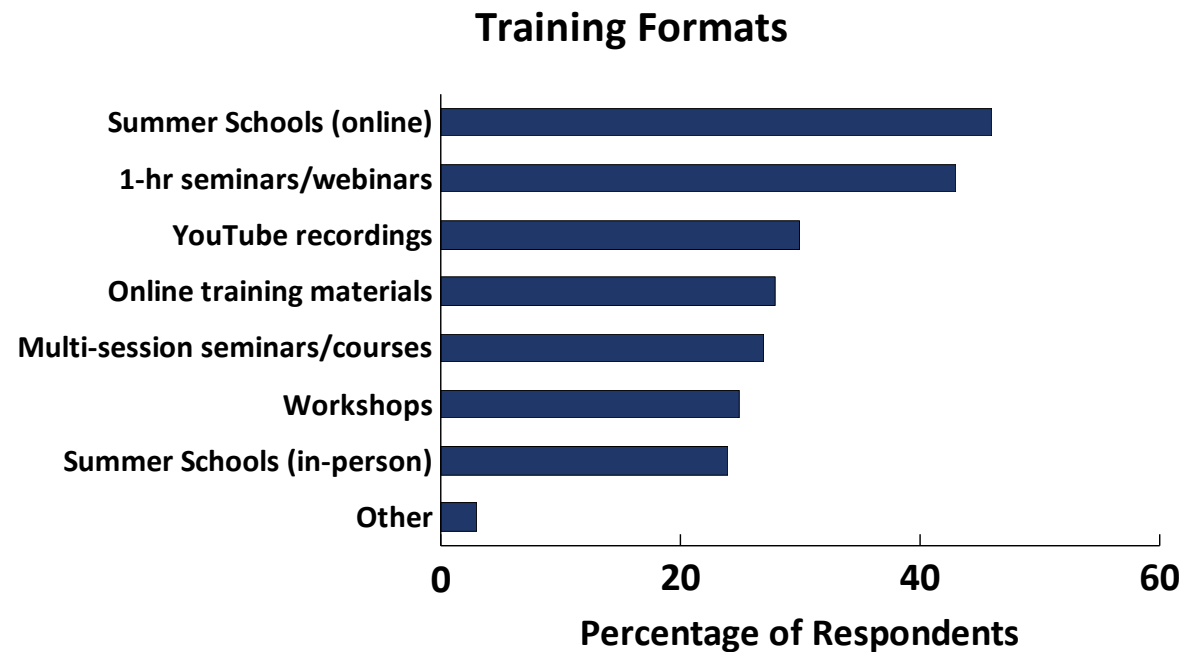


14% of respondents reported that they have received training from a different source.

n=85

Training Elements

Training Elements



46%, and 43% of respondents reported that they received training in the format of online summer schools, and seminars/webinars, respectively.

62% of respondents reported that they received training in the format of summer schools (online or in-person).

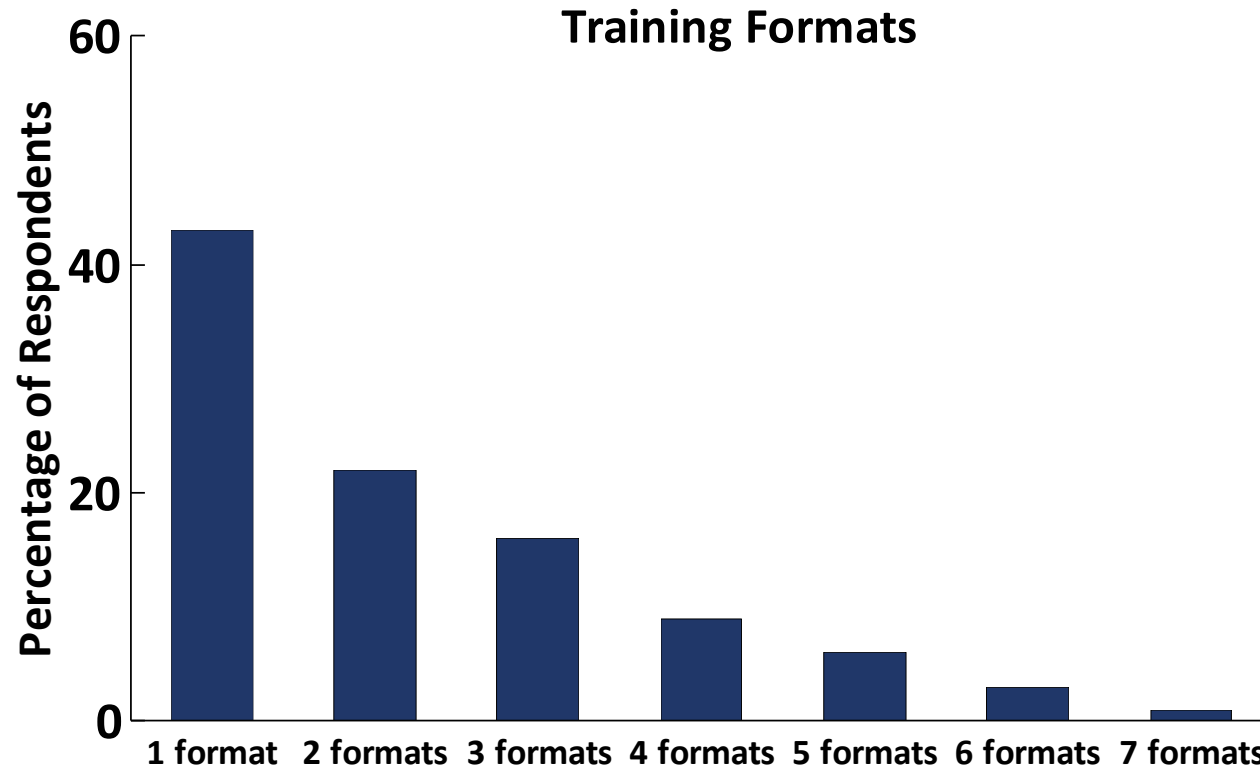
n=263

Prior to 2020, summer schools were in-person only however, due to COVID-19, they shifted exclusively to online learning for the past two summers (2020 and 2021). All other training sessions have also been exclusively online since the spring of 2020.

62 respondents reported that they attended in-person summer schools, indicating their longer-term engagement with the ARC training.

Of the 65 respondents outside Ontario, 86% (56 out of 65) reported that they received training in the format of online sessions.

Training Elements



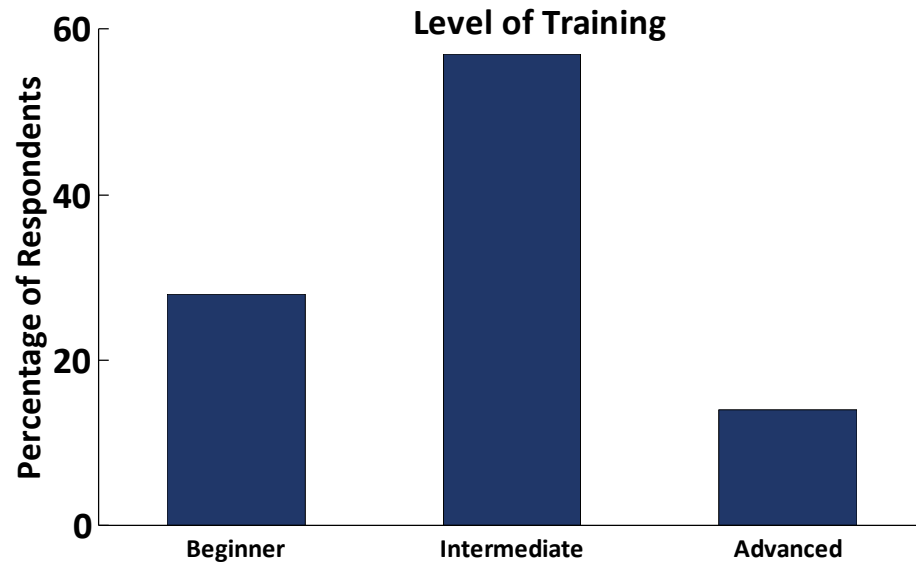
43% of respondents reported that they received training in only one format.

n=263

Most respondents reported that they received different types of training in multiple formats, specifically:

- **57% respondents reported that they received training in at least two different formats.**
- **35% respondents reported that they received training in at least three different formats.**

Training Elements



57% of respondents reported that their most recent training was at the intermediate level.

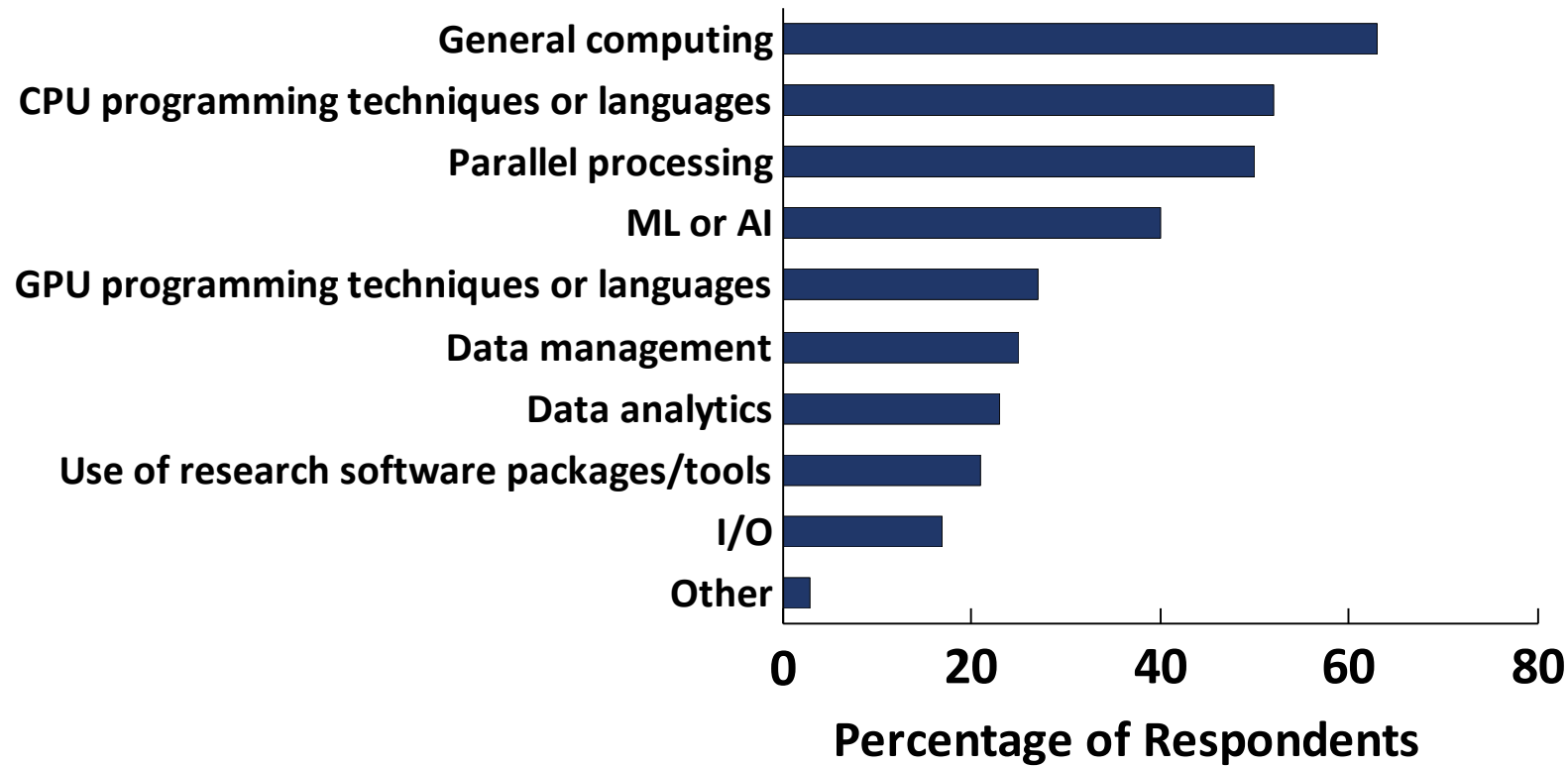
n=263

Faculty members, and postdoctoral fellows/researchers reported a relatively higher level of their most recent training compared to other students and researchers, specifically:

- 11% (29 out of 263) respondents are faculty members, of which 90% (26 out 29) reported that their most recent training was at the intermediate or advanced level.**
- 11% (28 out of 263) respondents are postdoctoral fellows/researchers, of which 82% (23 out 28) reported that their most recent training was at the intermediate or advanced level.**

Training Elements

Training Topics

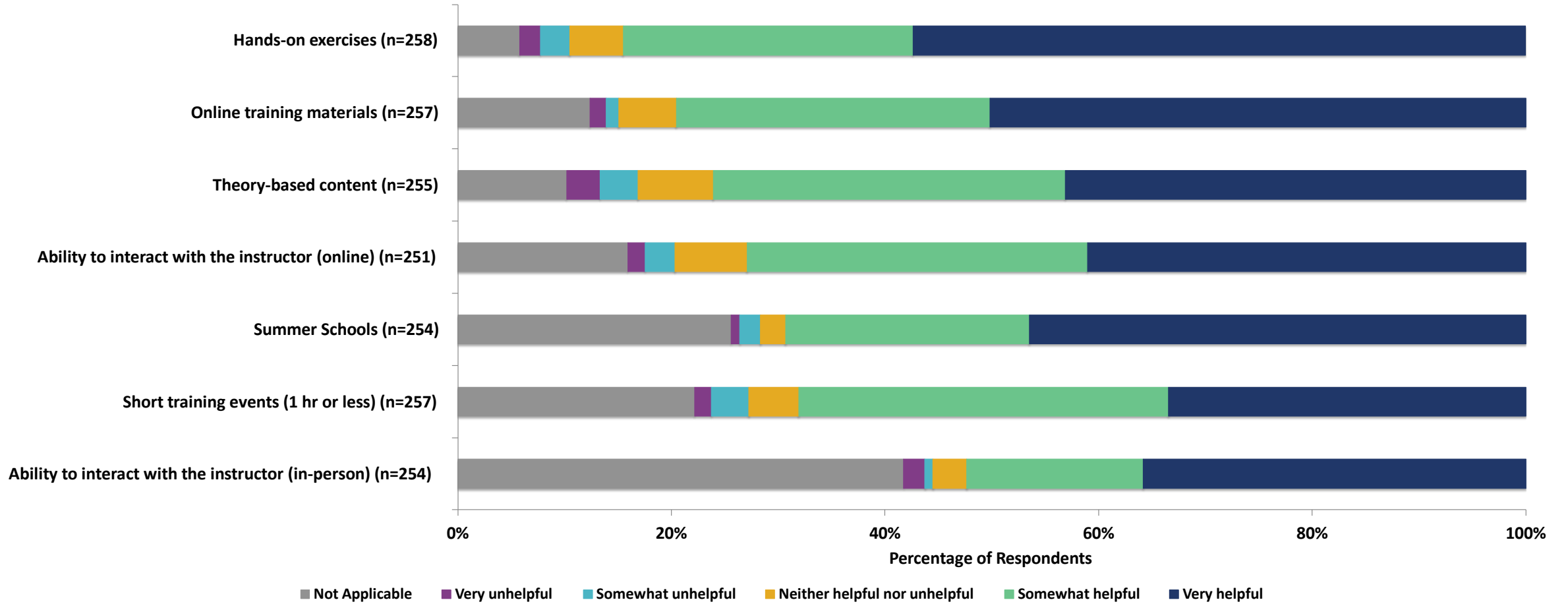


63% of respondents reported that general computing has been one of the topics covered in the training.

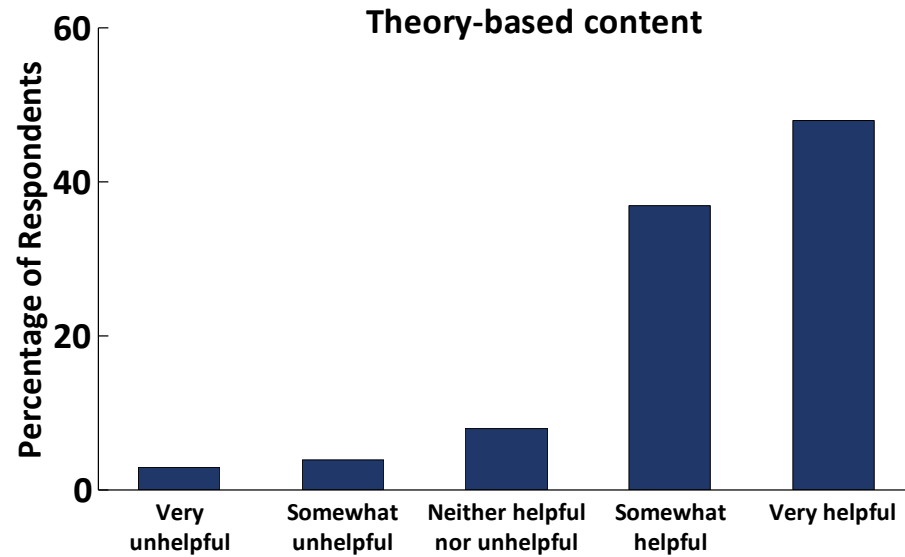
n=264

Researcher Perspectives on Training Elements

Researcher Perspectives

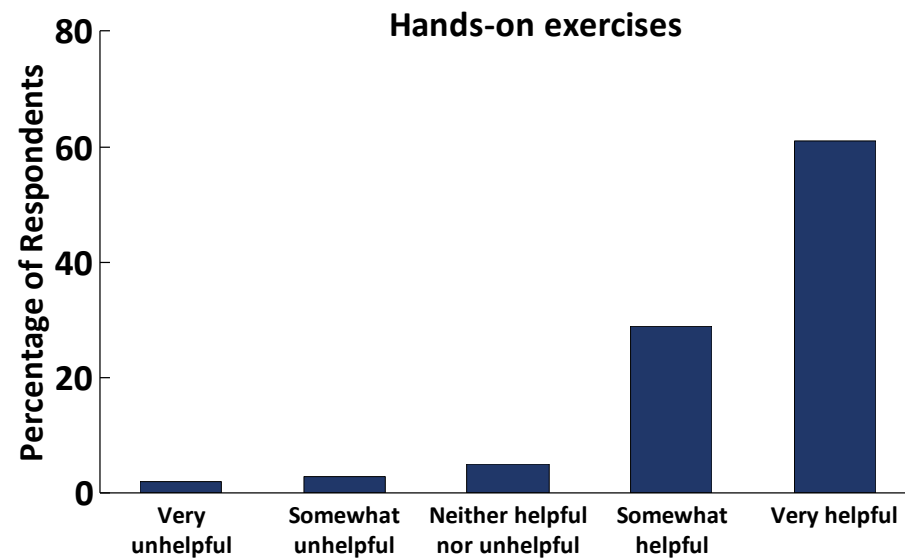


Researcher Perspectives



85% of respondents reported that they found the theory-based content to be helpful.

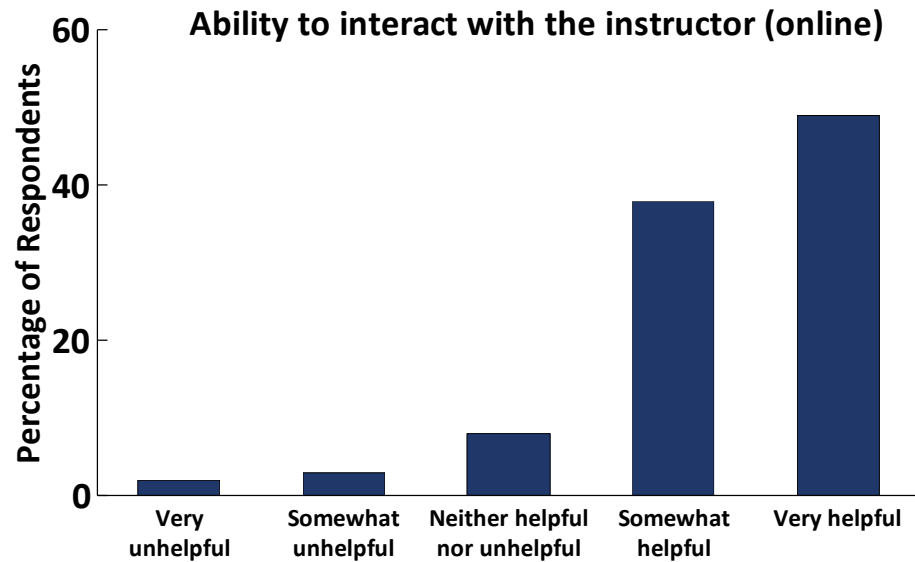
n=229



90% of respondents reported that they found the hands-on exercises to be helpful.

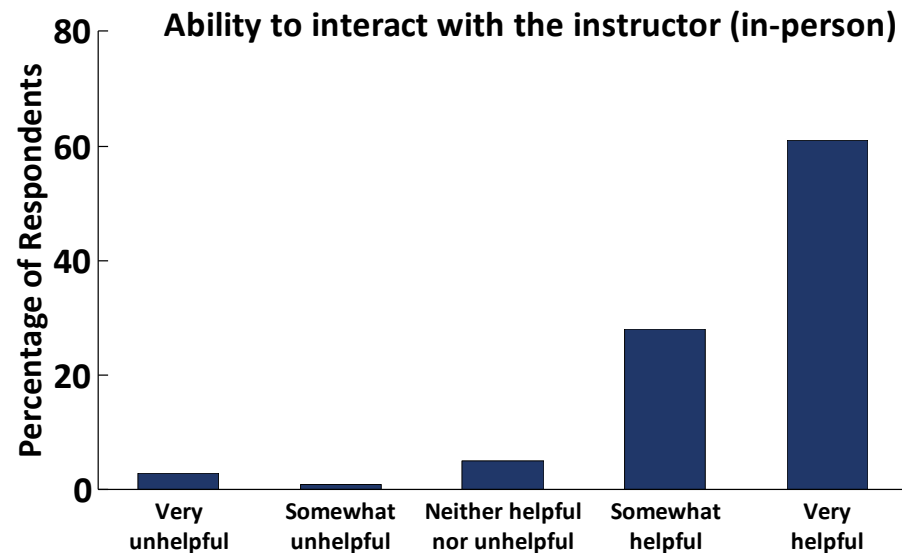
n=243

Researcher Perspectives



87% of respondents reported that they found the online interaction with instructor to be helpful.

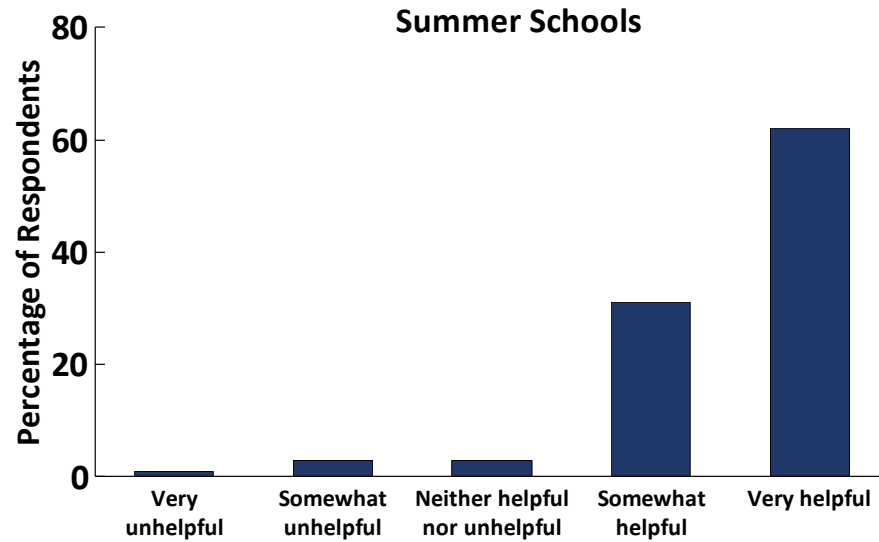
n=211



89% of respondents reported that they found the in-person interaction with instructor to be helpful.

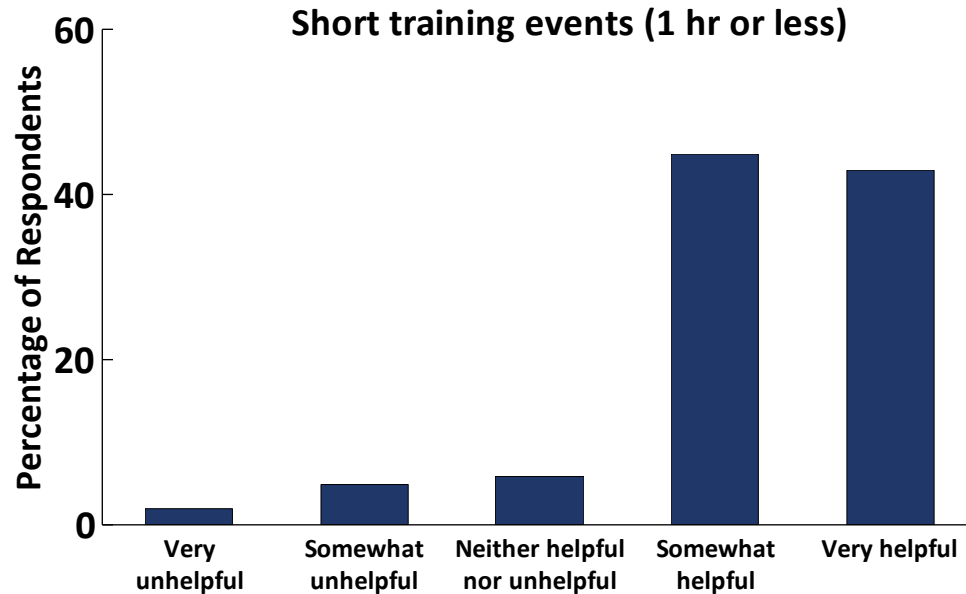
n=148

Researcher Perspectives



93% of respondents reported that they found the summer schools to be helpful.

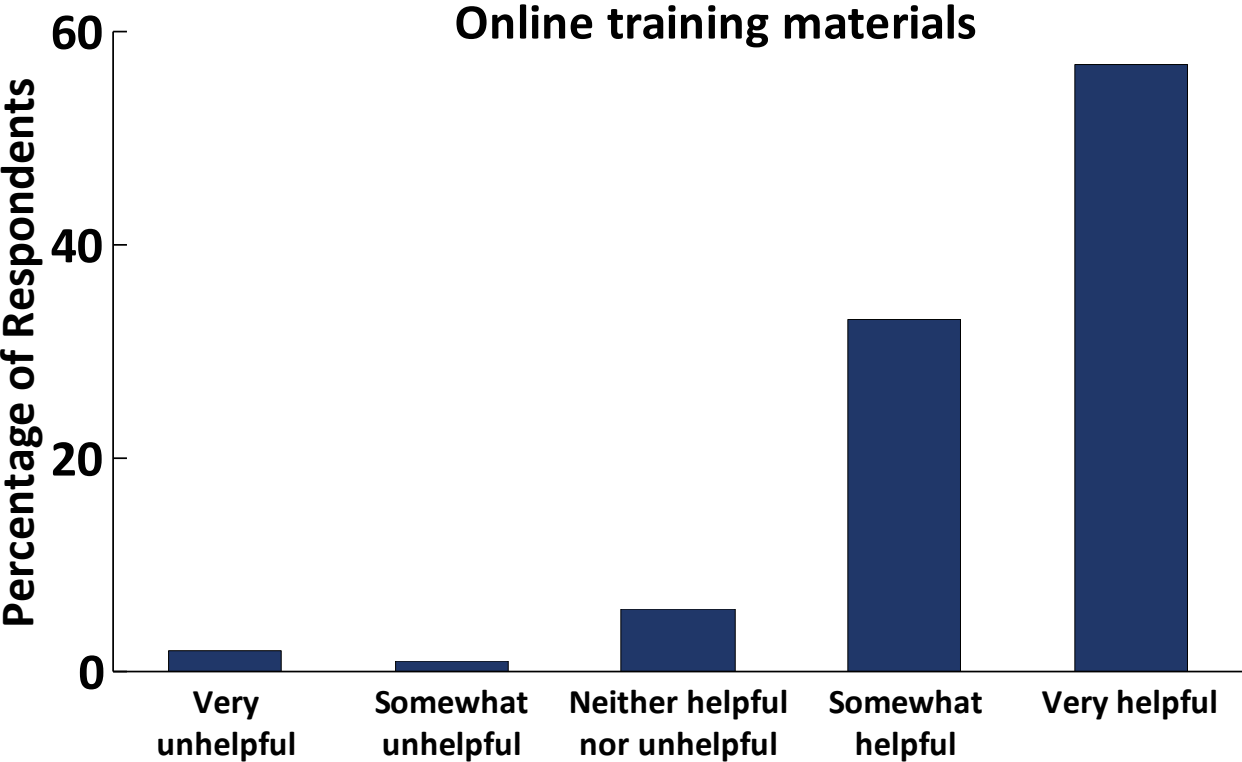
n=189



88% of respondents reported that they found the short training events to be helpful.

n=200

Researcher Perspectives

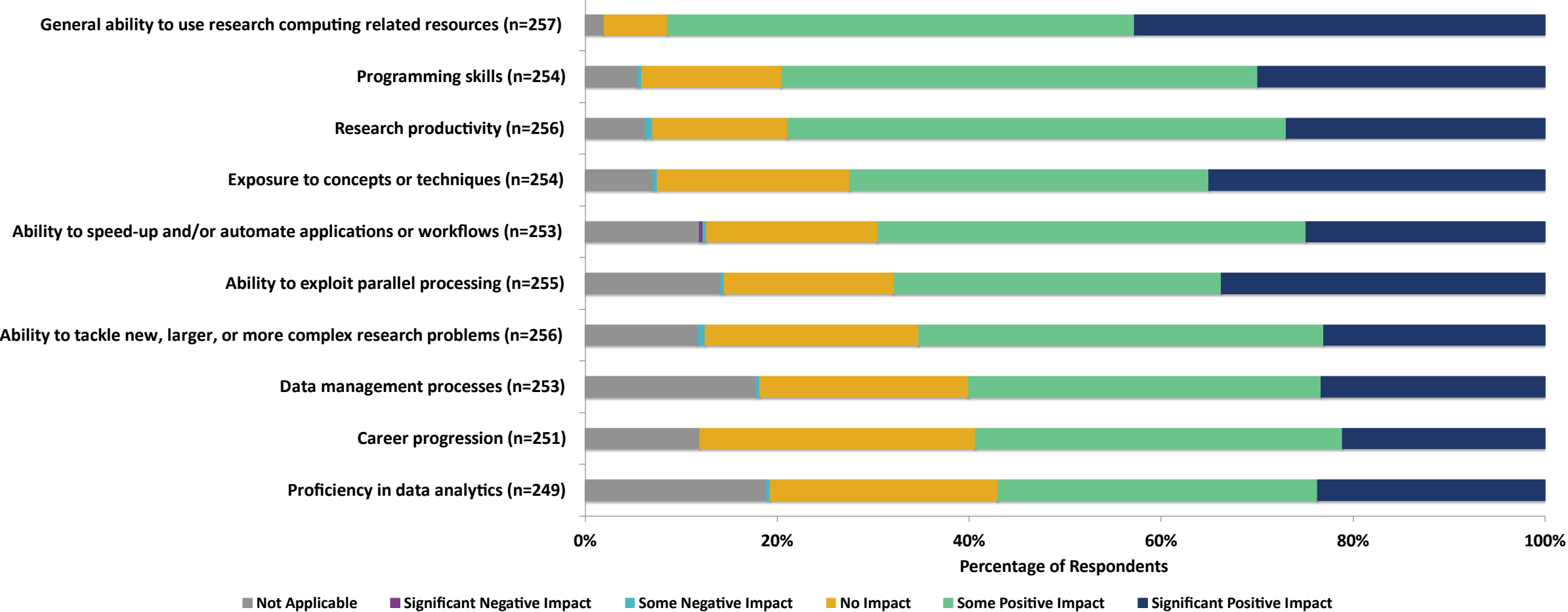


90% of respondents reported that they found the online training materials to be helpful.

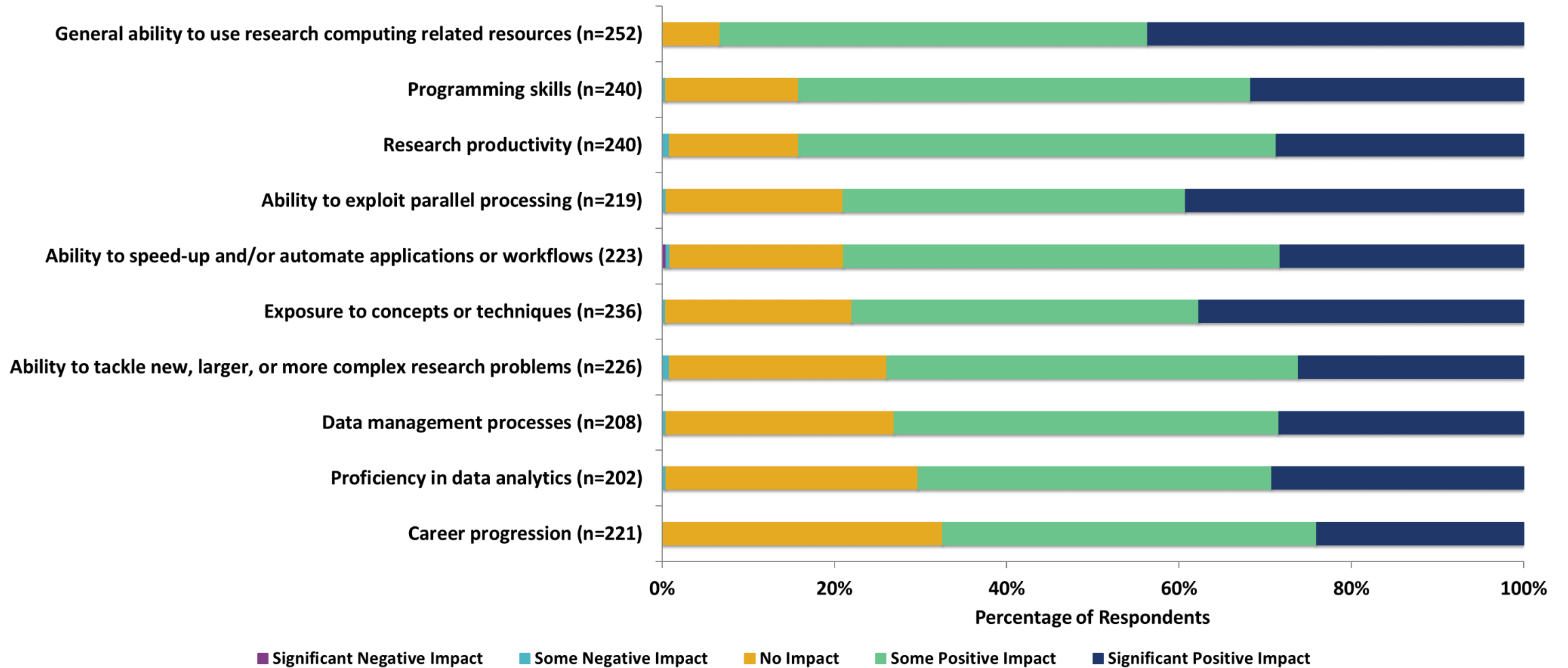
n=227

Impact of Training

Impact of Training

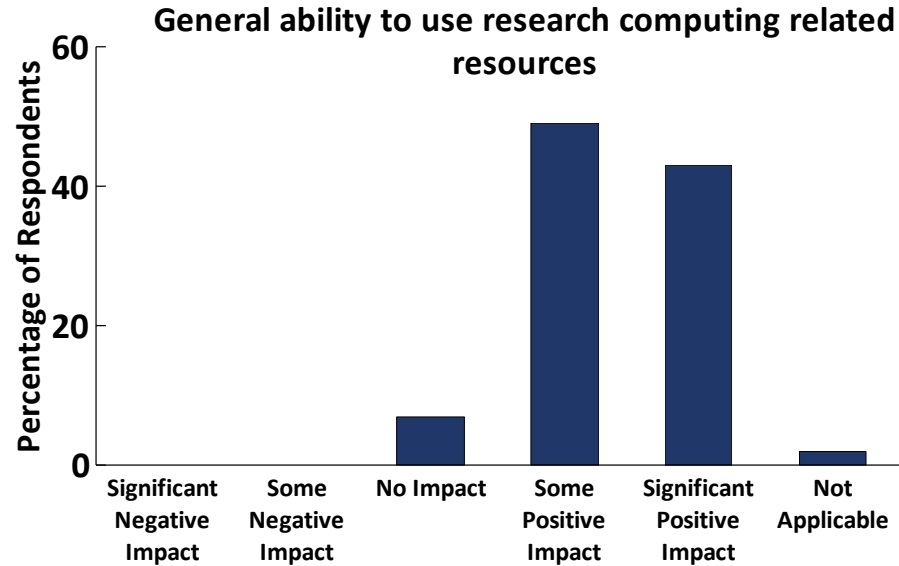


Impact of Training



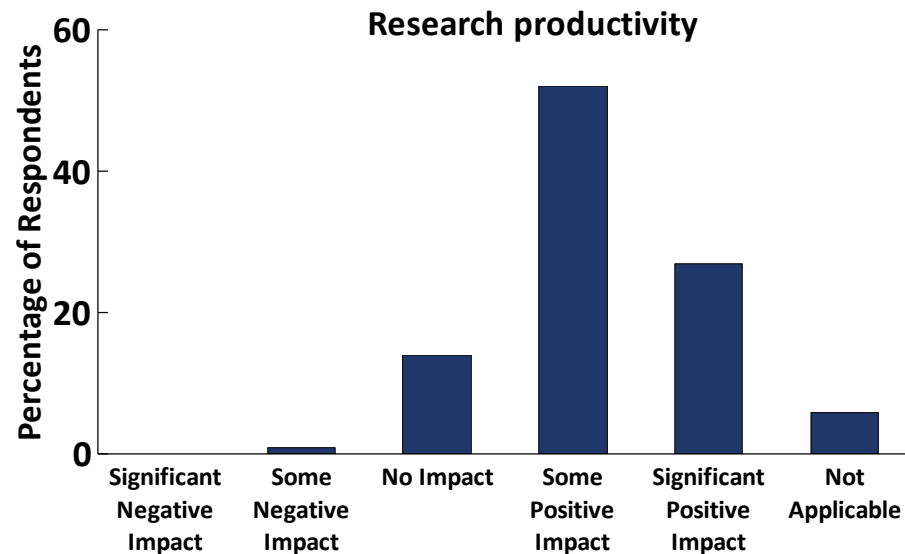
Note: 'Not applicable' responses are excluded.

Impact of Training



91% of all respondents reported positive impact of training on their general ability to use research computing related resources.

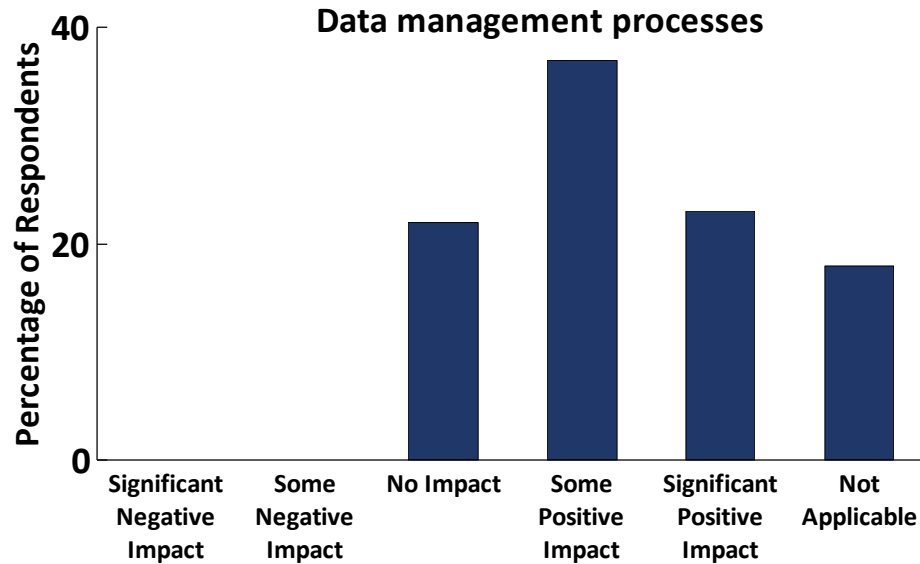
n=257



79% of all respondents reported positive impact of training on their research productivity.

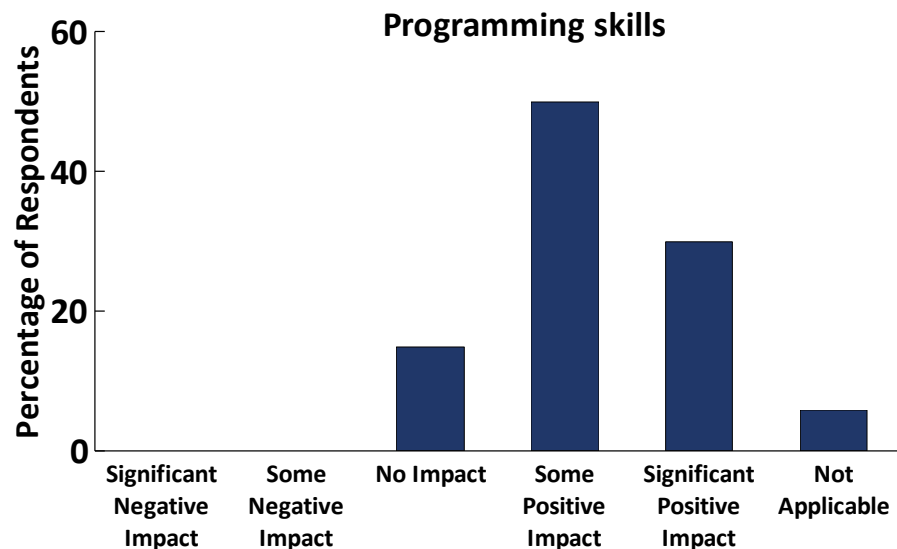
n=256

Impact of Training



60% of all respondents reported positive impact of training on data management processes.

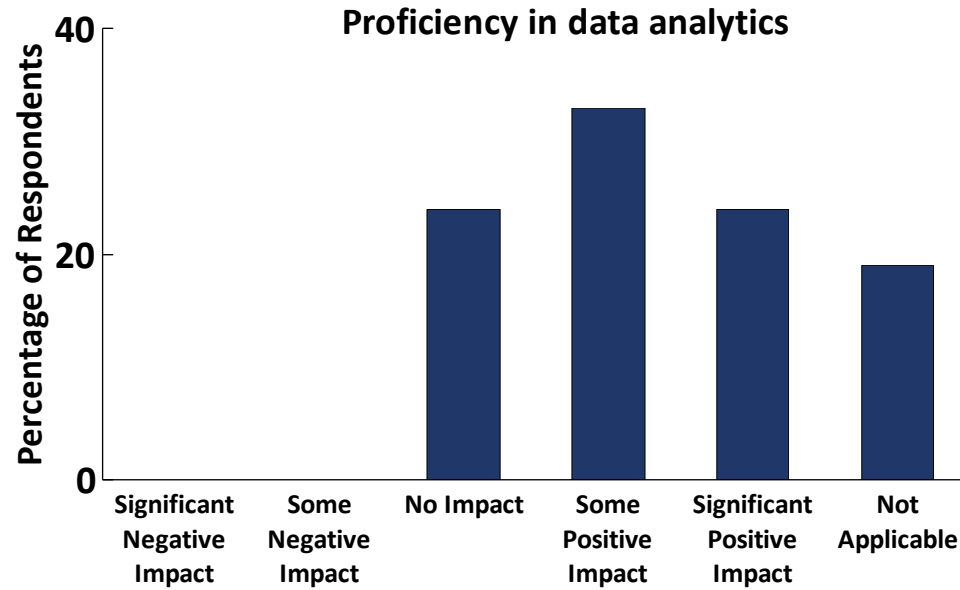
n=253



80% of all respondents reported positive impact of training on their programming skills.

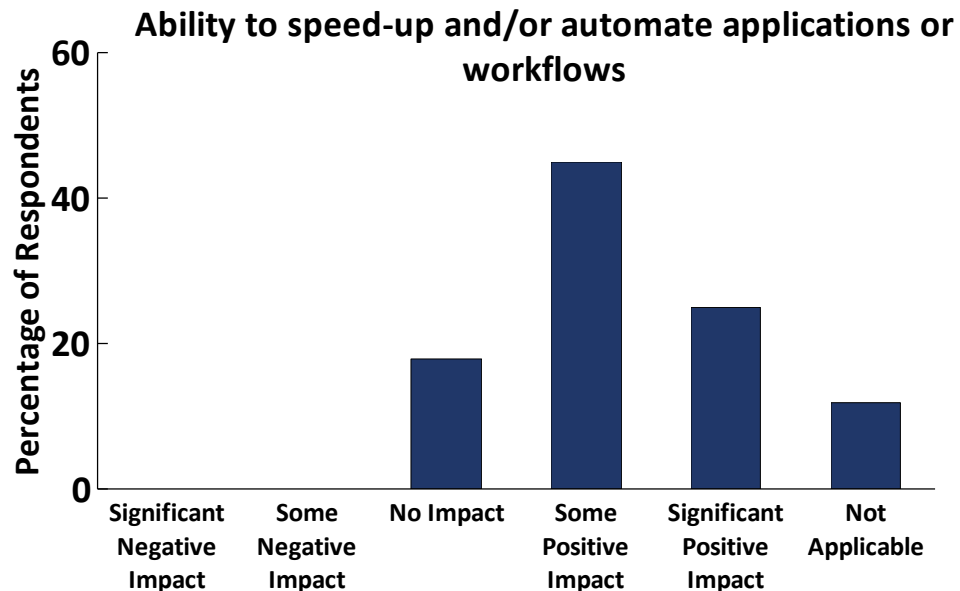
n=254

Impact of Training



57% of all respondents reported positive impact of training on their proficiency in data analytics.

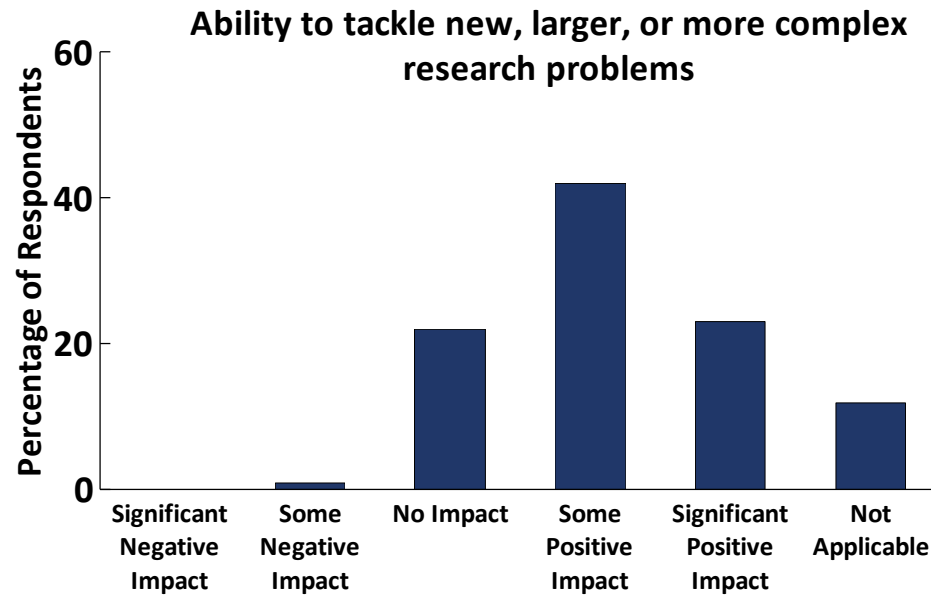
n=249



70% of all respondents reported positive impact of training on their ability to speed-up and/or automate applications or workflows.

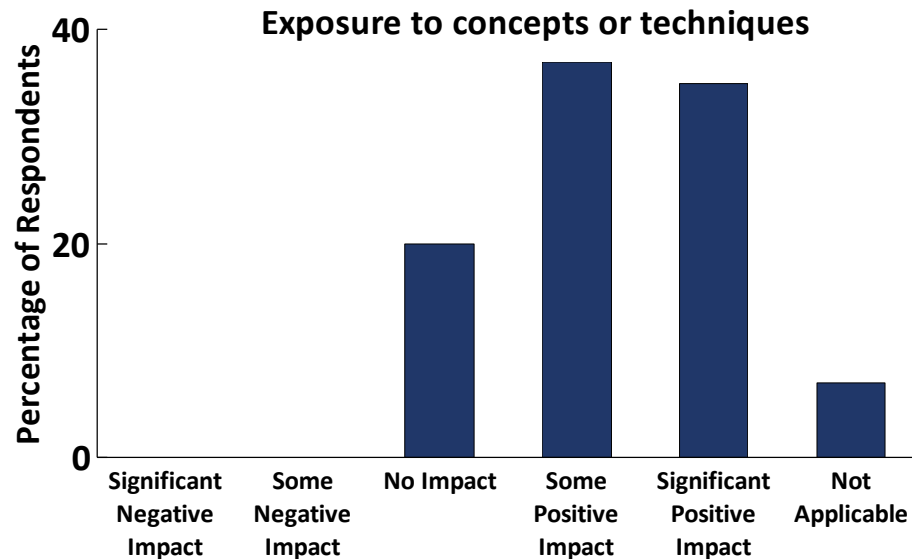
n=253

Impact of Training



65% of all respondents reported positive impact of training on their ability to tackle new, larger, or more complex research problems.

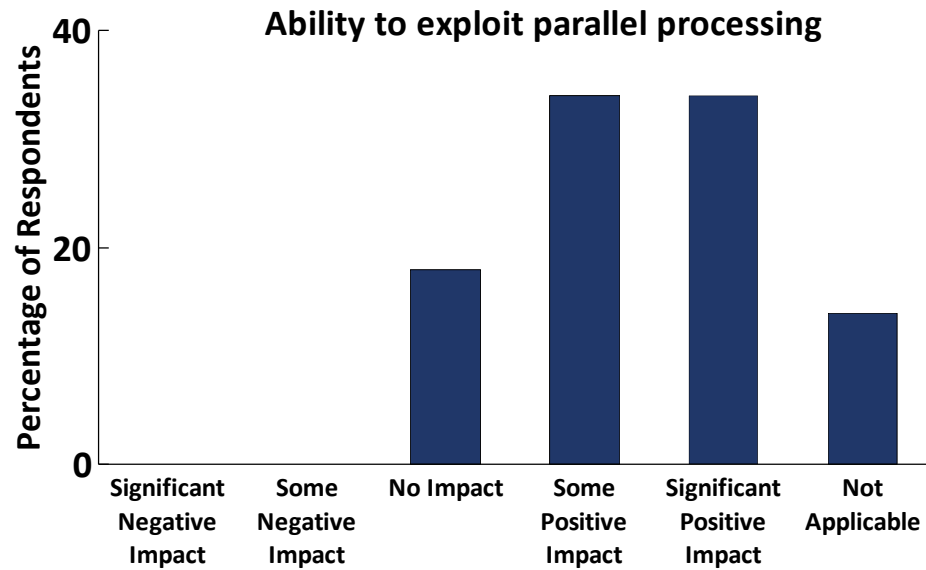
n=256



70% of all respondents reported positive impact of training on their exposure to concepts or techniques.

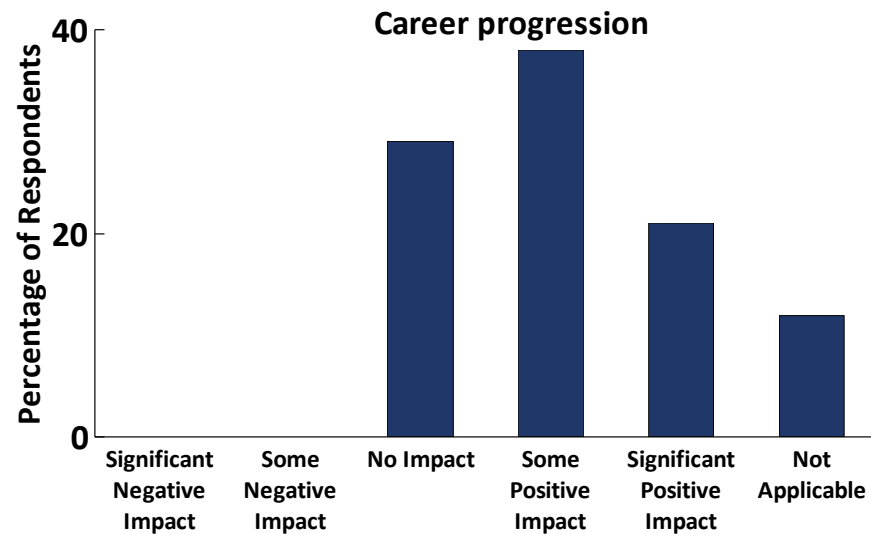
n=254

Impact of Training



68% of all respondents reported positive impact of training on their ability to exploit parallel processing.

n=256

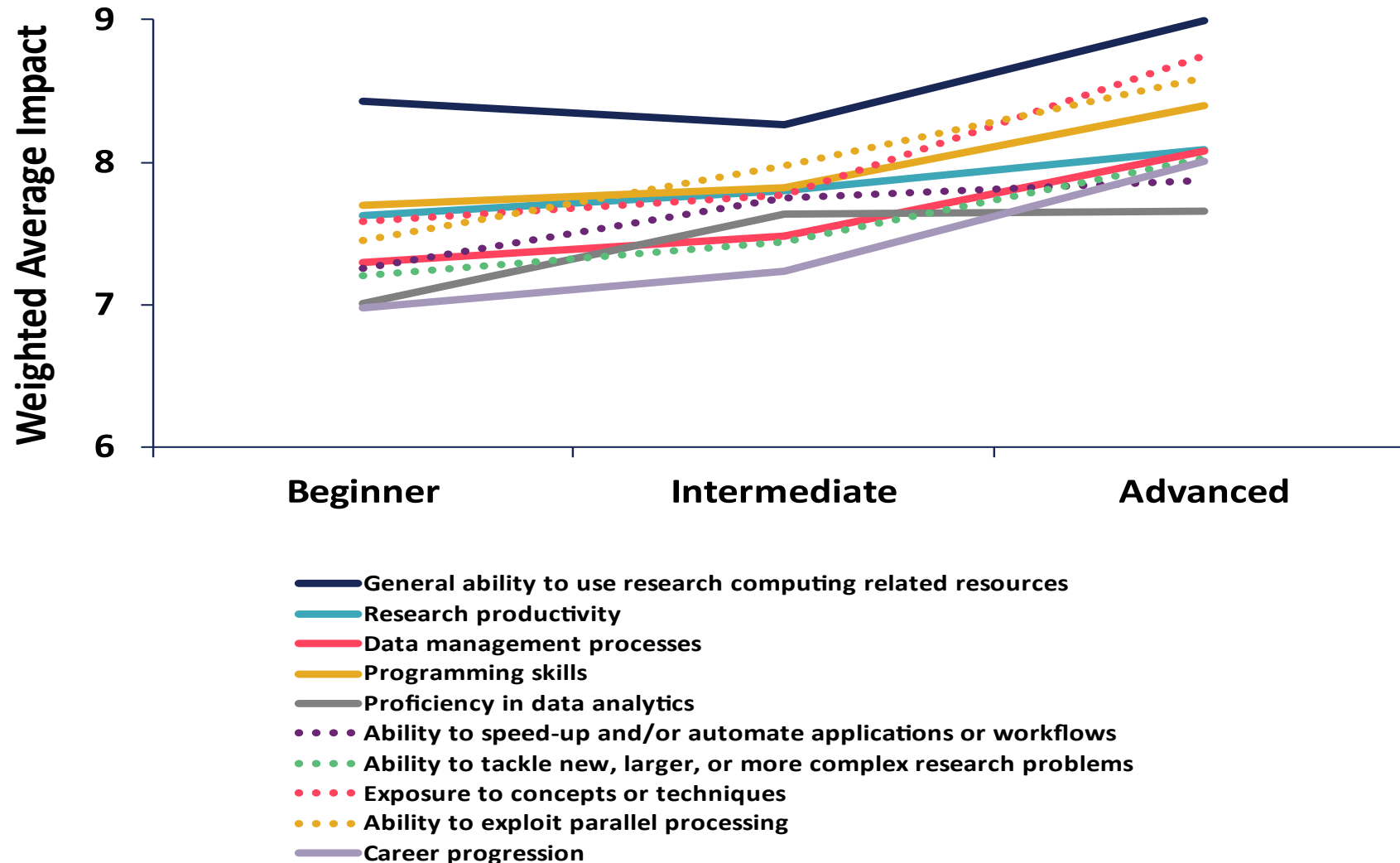


59% of all respondents reported positive impact of training on their career progression.

n=254

Impact of Training

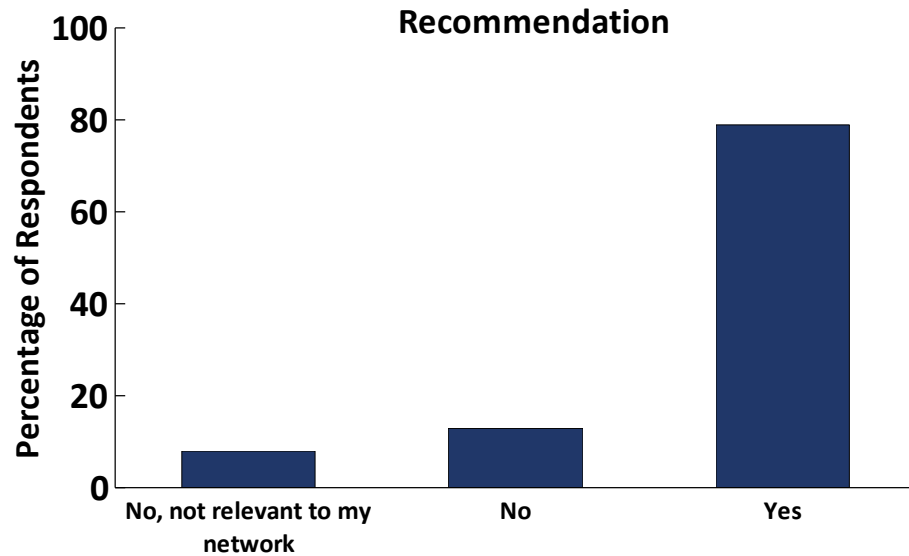
Average Impact on Research Capabilities and Career Progression
Segmented by Level of Training



In general, the average impact of the training on the capabilities and career progression of researchers is greater for those who received more advanced level of training, compared to those who received intermediate or beginner levels of training (significant at least at the 90% confidence level).

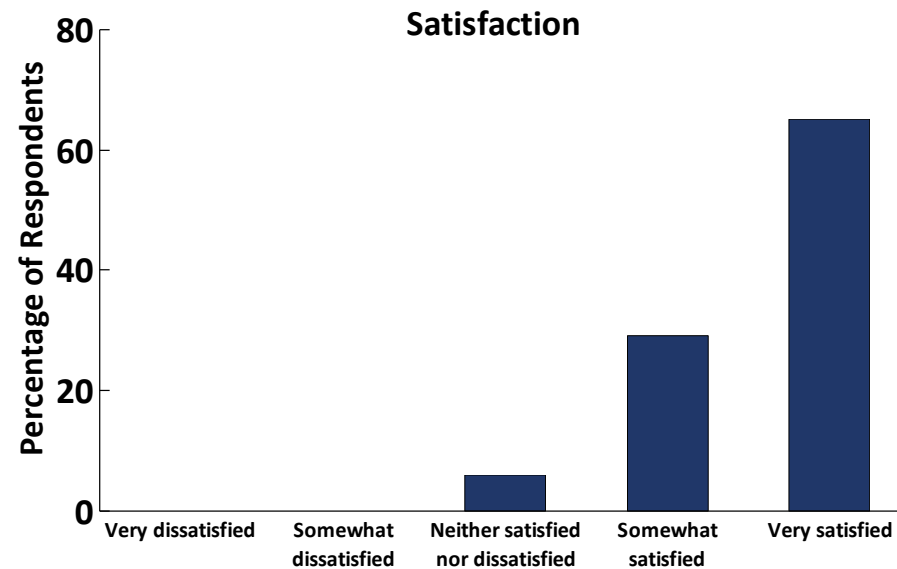
Overall Satisfaction

Overall Satisfaction



79% of respondents indicated that they have recommended the training to others in their network.

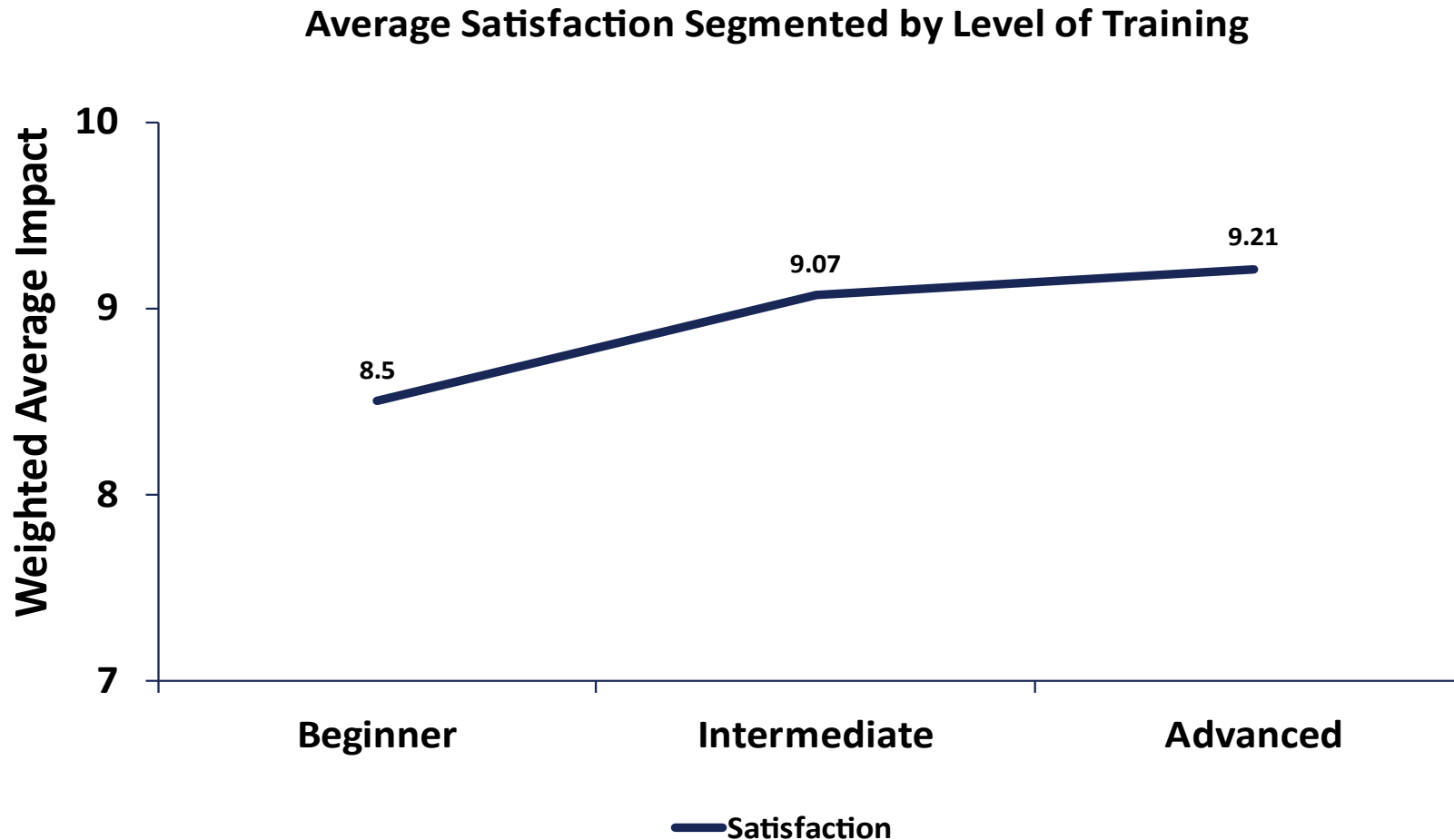
n=261



93% of respondents reported satisfaction with the training provided by the Ontario ARC consortia.

n=257

Overall Satisfaction



In general, the average satisfaction reported by researchers is greater for those who received more advanced level of training, compared to those who received intermediate or beginner levels of training (significant at the 95% confidence level).

Variables Associated with Impact

- Correlation analysis indicated that greater impacts on research capabilities and career progression have been reported by:
 - Researchers that attended a greater number of training sessions, and
 - Researchers that have recently received more advanced level of training

Additional Comments from Survey Respondents

Additional Training Topics Suggested by Respondents

- Access to HPC system
- Advanced bash scripting
- Advanced data analysis and statistical modelling, predictive modeling, exploratory data analysis, and explainable AI
- Basic bioinformatics for plant scientists
- Best practices for plotting
- C++, Java, SQL, Python, Fortran, MPI, OpenFOAM
- Code optimization and debugging
- Continuous development, integration
- Deploying VMs from Graham
- Distributed learning
- Debugging a network using TensorBoard
- Domain-specific packages (e.g., programs used for chemistry)
- Finite element software training (e.g., LS-Dyna, Abaqus, ANSYS)
- Functional programming, use of new and upcoming languages like Julia on HPC systems
- Interaction between cloud facilities at CC and local resources via REST, GRAPHQL or gRPC
- Julia simulation courses
- Machine learning with concrete examples of usage
- Multi-access Edge Computing (MEC) & 5G
- Seismological processing using parallel processing
- Set up a cloud computing environment from scratch, and how to use Docker
- Tensorflow and PyTorch training
- Validation first approach

Other Suggestions from Respondents

- More practical sessions for practice and demonstration of questions and answers
- More visualization workshops
- More introductory courses
- More shorter lectures and hands-on work
 - “More shorter lectures and hands-on work (a format adopted in later summer school courses) would be even more helpful.”
 - “I like short regular training on a hot topic every week like only in half an hour! Every topic can be explained in short even with small piece of code!”
- Training for data scientists to prepare them for job market
 - “...basic of software engineering such as testing, debugging, documentation, modular programming, packing, OOP, version controlling, refactoring, they can be more trained for job market.”
- Specify course requirements
 - “Some workshops require basic preexisting knowledge. It would be good to specify these requirements in the workshop description, so that participants take a full advantage of the instructors' time.”

Additional Comments from Respondents

Research capabilities and productivity

- “Overall, the summer school had left a positive impact on my programming skills. It helped me to accelerate my progress with more knowledge (for my own PhD research)...the summer school enhanced my abilities significantly.”
- “Provide opportunity to researchers to present their research, as a special seminar, at different platforms nationally, internationally and in colleges targeting student audience.”
- “The training I received allowed me to analyze my research data with much more depth. I am no longer slowed down by large datasets or the limits of MS Excel.”
- “The training provided by ARC benefits my students directly. Therefore, it helps my group's research productivity and I do not have to find other ways for my students to get this training (which is what I had to do many years in the past).”
- “The training has helped improve my productivity in my PhD research.”
- “The training has helped me to redevelop my research focus and apply for more funding options.”

Additional Comments from Respondents

Career development and job market

- “The training was critical to my graduate school projects, then my postdoctoral, and now junior faculty projects as well as my career growth. I would not have started exploring increasingly complex modeling projects without the confidence and skills development (and hands-on training exercises) that the SciNet courses and summer schools provided.”
- “The training I received kickstarted my career in computational biology and helped me gain confidence in the field -- more than any other online/in-person training I received.”
- “In my field (neuroscience), data science and deep learning are hot issue, so I would like to get the training to improve and develop my career.”
- “I am currently working in financial investments and there has been a shift to automation and big data in recent years where I am now using python that I learned from SciNet.”

Additional Comments from Respondents

Responsive staff and instructors

- “In particular, having day-long interactive workshops in which hands-on coding examples could be tackled with the instructors present to answer questions was extremely beneficial. Not only did this provide working examples that we could refer to after the class ended but the experience of coding it for oneself provided a deeper, longer lasting understanding of the material.”
- “Dr. Marcello Ponce from SciNet has helped me through my graduate training more than my supervisor and immediate colleagues, both in terms of encouragement and problem solving.”
- “I really appreciate that the instructors make themselves available to answer questions beyond the seminars & courses.”
- “I have found personnel from SHARCNET to be competent and responsive.”

Introductory courses and basic instructions

- “Many of the training sessions assume people know the capabilities of software and how it can be applied. Some more basic instruction on what the capacity of these systems are would be very useful.”
- “The courses were very advanced, and I think there should be some prerequisite courses.”

Theoretical part was not that useful

- “...theoretical part in some training was not useful. for example, for the short session for ssh key, there was a lot of redundant stuff at the start of the session that just made it really hard to follow up what is the purpose of the session and I ended up checking it myself online.”