ASCR Workshop on Extreme Heterogeneity
January 23-25, 2018

Report to the Advanced Scientific Computing Advisory Committee

April 17, 2018

Lucy Nowell, PhD

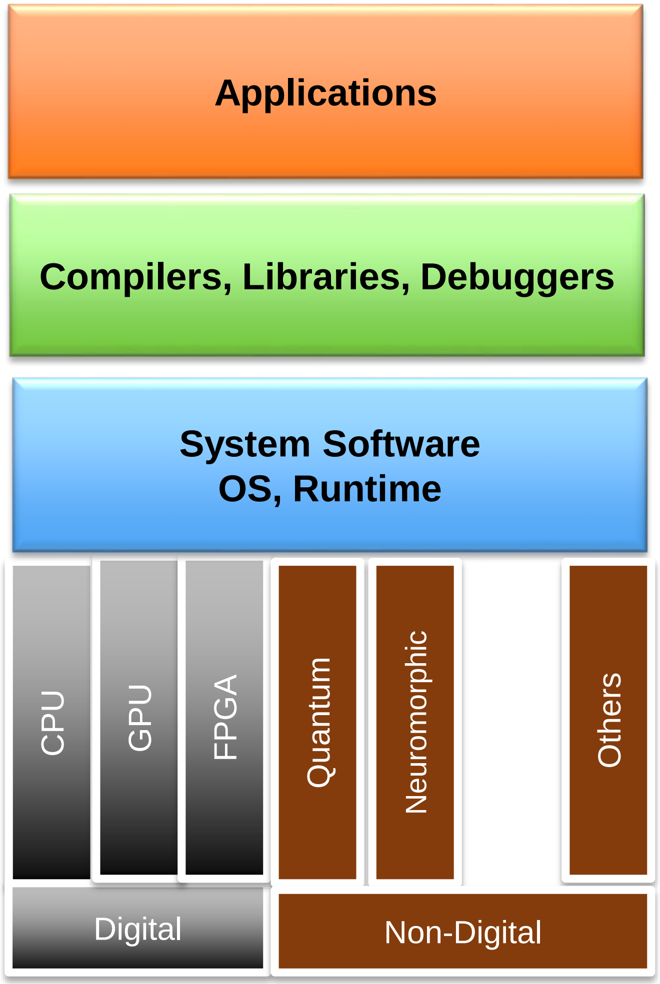
Computer Science Program Manager

Advanced Scientific Computing Research

ASCR Extreme Heterogeneity Workshop

January 23-25, 2018 Virtual Meeting

- **POC: Lucy Nowell (Lucy.Nowell@science.doe.gov)**
- **Goal: Identify Priority Research Directions for Computer Science needed to make future supercomputers usable, useful and secure for science applications in the 2025-2040 timeframe**
- Primary focus on the software stack and programming models/environments/tools.
- 150+ participants: DOE Labs, academia, & industry
- Observers from DOE and other federal agencies
- Factual Status Document (FSD)
- White papers solicited (106 received!) to contribute to the FSD, identify potential participants, and help refine the agenda
  - Draft Priority Research Directions delivered to ASCR AD Barb Helland on 3/5/18.
  - Report due early May 2018



Workshop Organizers & Program Committee

- Jeff Vetter (ORNL), Lead Organizer and Program Committee Chair
- Organizing Committee and Program Committee Members: Ron Brightwell (Sandia-NM), Pat McCormick (LANL), Rob Ross (ANL), John Shalf (LBNL)
- Program Committee Members: Katy Antypas (LBNL, NERSC), David Donofrio (LBNL), Maya Gokhale (LLNL), Travis Humble (ORNL), Catherine Schuman (ORNL), Bryan VanEssen (LLNL), Shinjae Yoo (BNL)

Workshop Statistics

- DOE National Lab Participants
  - Invited: 85, representing 11 Labs
  - Registered: 82
- Academic Participants
  - Invited: 42, representing 38 institutions
  - Registered: 32
- Industry Participants
  - Invited: 14, representing 11 companies
  - Registered: 11

Workshop Charge Excerpts

The purpose of this workshop is to identify the priority research directions for ASCR in providing a smart software stack that includes techniques, such as deep learning to make future computers composed of a variety of complex processors, new interconnects and deep memory hierarchies easily used by a broad community of computational scientists…The primary aim for the workshop is to identify the new algorithms and software tools needed from basic research in computer science to enable ASCR's supercomputing facilities to support future scientific and technological advances on SC program's grand challenge problems. ASCR's grand challenges and the resulting priority basic research directions should be identified by spanning existing and next generation computer architectures, including novel technologies that may be developed in the "Post-Moore's Law era" and the promising tools and techniques that are essential to efficient and productive utilization of such architectures. The workshop and subsequent report should define basic research needs and opportunities in computer science research to develop smart and trainable operating and runtime systems, execution models, and programming environments that will make future systems easier to tailor to scientists' computing needs and for facilities to securely deploy.

Workshop Plenary Talks

- Welcome and ASCR Update – Barbara Helland, Director, Advanced Scientific Computing Research
- View from ASCR Research Division – Steve Lee, Acting Division Director
- Invited Plenary Talk: IEEE Rebooting Computing - Tom Conte
- Architectural Trends and System Design Issues - Bob Colwell
- Introduction to Extreme Heterogeneity – Jeffrey Vetter, John Shalf, and Maya Gokhale + Factual Status Document section leads
- Report on the ASCAC Future of Computing study - Maya Gokhale
- Panel on Issues Raised by Extreme Heterogeneity - Moderator Ron Brightwell
  - Usability, Understandability and Programmability – Salman Habib
  - Operating and Runtime Systems – Ron Brightwell
  - Data Analytics – Wes Bethel
  - EH Workflow Management – Ewa Deelman
- Memory Systems and I/O - Bruce Jacob
- ORNL/DoD Beyond CMOS Workshops - Neena Imam
- Exascale Computing Project Computer Science R&D for Heterogeneity – Mike Heroux

Breakout Groups and Moderators

1A - Programming Environments, Models, & Languages – Alex Aiken & Pat McCormick

1B - Data Management and I/O – Moderators Rob Ross & Suren Byna

1C - Data Analytics and Workflows – Tom Peterka & SJ Yoo

2A - Operating Systems and Resource Management – Ron Brightwell

2B- Software Development Methodologies – Sherry Li

2C - Modeling & Simulation for Hardware Characterization – Andrew Chien & David Donofrio

3A - Programming Environments: Compilers, Libraries and Runtimes – Michelle Strout & Barbara Chapman

3B - Data Analytics and Workflows – Christine Sweeney

3C - System Management, Administration & Job Scheduling –Paul Peltz & Rebecca Hartman-Baker

3D - Crosscut: Productivity, Composability, & Interoperability – Bob Lucas

4A - Programming Environments: Abstractions, Models and Languages – Pat McCormick

4B - Crosscut: Modeling and Simulation – Jeremy Wilke & Zhiling Lan

4C - Operating Systems & Resource Management: Locality & Programming Environment Support - Mike Lang

4D -Crosscut: Portability, Code Reuse and Performance Portability - Anshu Dubey

5A -Data Management and I/O – Rob Ross

5B- Programming Environments: Debugging, Autotuning, & Specialization – Mary Hall & John Mellor-Crummey

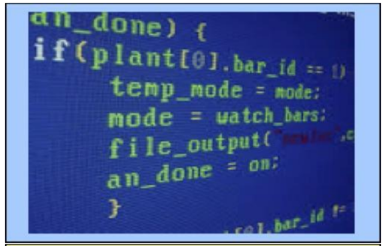
5C -Crosscut: Resilience & Power Management –Franck Cappello & Kirk Cameron

What Do We Mean by Extreme Heterogeneity?

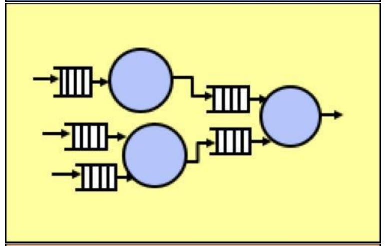
- **Exponentially Increasing Parallelism** (central challenge for ECP, but will be even worse)
    - **Trend**: *End of exponential clock frequency scaling (end of Dennard scaling)*
    - **Consequence:** *Exponentially increasing parallelism*
- **End of Lithography as Primary Driver for Technology Improvements**
    - **Trend:** *Tapering of lithography Scaling*
    - **Consequence:** *Many forms of heterogeneous acceleration (not just GPGPUs anymore)*
- **Data Movement Heterogeneity and Increasingly Hierarchical Machine Model**
    - **Trend**: *Moving data operands costs more than computation performed on them*
    - **Consequence**: *More heterogeneity in data movement performance and energy cost*
- **Performance Heterogeneity**
    - **Trend**: *Heterogeneous execution rates from contention and aggressive power management*
    - **Consequence**: *Extreme variability and heterogeneity in execution rates*
- **Diversity of Emerging Memory and Storage Technologies**
    - **Trend**: *Emerging memory technologies and stall in disk performance improvements*
    - **Consequence**: *Disruptive changes to our storage environment*
- **Increasingly Diverse User Requirements**
    - **Trend**: *Diverse and Complex and heterogeneous scientific workflows*
    - **Consequence**: *Complex mapping of heterogeneous workflows on heterogeneous systems.*

Extreme Heterogeneity PRDs

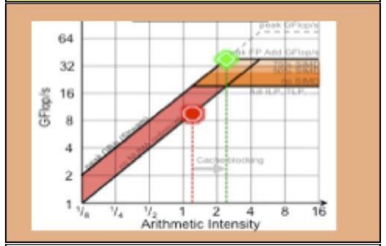
Reduce time to verifiable discovery despite diverse application domains and an exponential increase in architectural complexity from rapidly changing heterogeneous systems:



- **Maintaining and improving programmer productivity**
  - Flexible, expressive, programming models and languages
  - Intelligent, domain-aware compilers and tools
  - Composition of disparate software components



- **Managing resources intelligently**
  - Automated methods using introspection and machine learning
  - Optimize for performance, energy efficiency, and availability

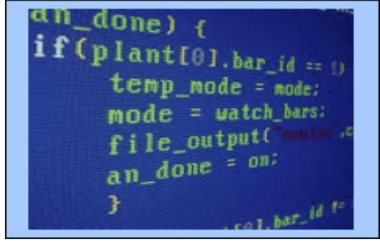


- **Modeling & predicting performance**
  - Evaluate impact of potential system designs and application mappings
  - Model-automated optimization of applications

- **Enabling reproducible science despite non-determinism & asynchronicity**
  - Methods for validation on non-deterministic architectures
  - Detection and mitigation of pervasive faults and errors

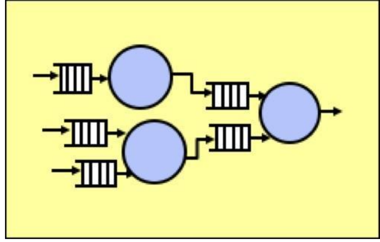
- **Facilitating Data Management, Analytics, and Workflows**
  - Mapping of science workflows to heterogeneous hardware & software services
  - Adapting workflows & services to meet facility-level objectives through learning approaches

Maintaining & Improving Programming Environments



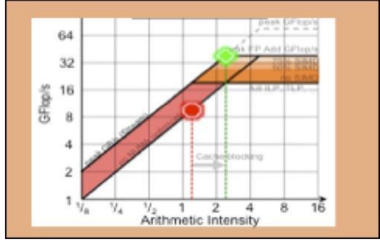
- **Programmability:** A diverse range of applications and memory/storage and hardware accelerator technologies will dramatically raise the complexity of programming challenges. Our programming mechanisms must change to maintain, let alone improve, the *current time-to-discovery*.
  - Requires more sophisticated implementations and range of capabilities (descriptive and prescriptive techniques) to fully utilize system capabilities.
  - Must be combined with higher levels of abstraction and domain-awareness to reduce complexity for user-facing interfaces, and increase flexibility and introspection for low-level developer-facing interfaces.
- **Intelligent toolchains:** Must assist and enable manual and/or automatic (ML/AI-based) selection of: resources for execution, data movement and transformations, and dynamic optimization for complex applications and workflows.
  - This requires more advanced compilers and supporting technology infrastructure.
  - Better tools for understanding performance profiling, debugging, and the associated state of the system and its components (introspection and real time machine learning mechanisms).
  - Runtime systems must support "smart" scheduling and affinity across in dynamic and diverse application and system architecture environments.
- **Composition**: Software reuse and interoperability will be key to controlling *costs and accelerating time-to-discovery.*
  - The software stack - from languages, tools, and OS/R components - must support effective and efficient composition and execution of different high-level software components and across increasingly heterogeneous hardware.
  - Efforts at the lower-levels of the stack will become key building blocks for coordination and execution of higher level features and domain-aware abstractions.

Managing Resources Intelligently



- **OS/RTS Design:** Hardware resources will become more complex and diverse. The operating system (OS) and runtime system (RTS) must integrate special-purpose devices and accelerators. The OS cannot assume all resources on a node are identical and dedicated devices
  - OS/RTS must be efficient and sustainable for an increasingly diverse set of hardware components
  - Must provide capability for dynamic discovery of resources as power/energy constraints impose restrictions on availability
- **Decentralized resource management:** New scalable methods of coordinating resources must be developed that allow policy decisions and mechanisms to co-exist throughout the system. Hardware resources are becoming inherently adaptive, making it increasingly complex to understand and evaluate optimal execution and utilization
  - System software must be enhanced to coordinate resources across multiple levels and disparate devices in the system
  - Must leverage cohesive integration of performance introspection and programming system abstractions to provide more adaptive execution
- **Autonomous resource optimization**: Responsibility for efficient use of resources must shift from the user to the system software; must employ sophisticated and intelligent approaches optimize selection of resources to application needs
  - Need more automated methods using machine learning to optimize the performance, energy efficiency, and availability of resources for integrated application workflows
  - More sophisticated usage models beyond batch-scheduled, spaced-shared nodes adds significant complexity to the management of system resources

Modeling & Predicting Performance



Understand, predict, & steer application behavior in EH environments

- **Develop methodologies and tools for accurate and fast modeling of EH at scale**
  - Accurately and quickly explore design space of heterogeneous function units, memories, and interconnects
  - Design methodologies and tools to understand system design tradeoffs in using heterogeneous components, including power, heat, resilience, and overall impact on applications
  - Provide integrated architecture and application modeling
    - Capture realistic workloads
    - Model diverse compute, storage, and interconnect technologies
  - Investigate new methods for model building and runtime cost reduction
    - Stochastic methods, and ML
    - Reduced precision and other approximation techniques
  - **Impact will be transformative**
    - From hardware components to workflow design
    - Predictive and accurate evaluation enabling optimization of new EH systems

- **Cohesively integrate modeling & simulation infrastructure with programming models/environment**
  - Support performance portability across many target EH architectures by providing cost models to aid both application developers and compilers/runtime
  - Invent new techniques to combine non-von (neuromorphic, quantum) with “more Moore”
  - Leverage ML and synthesis of mathematical models to tune compiler/runtime options from high-dimensional search space
  - **Near term impact to steer development/delivery of EH platforms and applications**
  - **Long term impact to workflow and programming environments guided by modsim cost models**

Enabling Reproducible Science

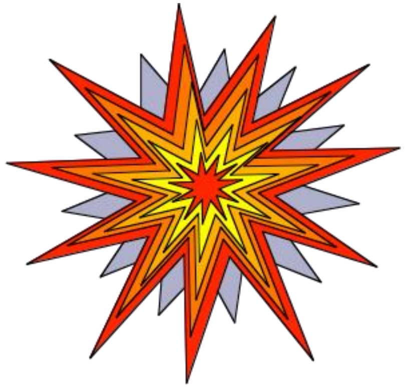
- **Reproducibility: Determination of verifiable computational results faces significant challenges in an increasingly diverse and heterogeneous environment**
  - Different types of processors/accelerators will not produce identical results for the same computations. Past techniques such as bitwise reproducibility are no longer applicable. These issues will be amplified by dynamic application behaviors, asynchronous and non-deterministic execution, and differences in capabilities across systems.
  - Additional challenges arise from non-Von Neumann architectures (e.g. quantum and neuromorphic) and machine learning techniques that can introduce “approximate” solutions.
- **Faults and Errors: The detection and mitigation of pervasive faults and errors becomes more complex with different behaviors and design details of an increasing number of hardware components**
  - Verification at different granularities becomes significant when reproducibility may not be possible or too expensive.
  - Must find ways to build effective and efficient testing strategies in the face of a combinatorial explosion of different realizations and compositions of heterogeneous processors
  - Must develop a more robust understanding of how to evaluate the validity of code and results.

Facilitating Data Management, Analysis, and Workflows

- **Workflow execution:** EH systems and applications bring significant challenges to usability through an unprecedented variety and number of data, resources, and services.
  - Discovery and mapping of science workflow requirements to appropriate data, hardware, and software services
  - New methods of composing scientific workflows (e.g., via workflow motifs)
  - Interfaces that facilitate HW and service composition – programming environments for data services
  - **Ties to Programmer Productivity**
- **Autonomous workflows:** Extracting the highest value from EH systems requires rapid, complex balancing across a wide variety of storage and networking technologies in response to a widely varying workload and significant rate of faults.
  - Profiling workflow telemetric data and learn what and how to instrument these systems
  - Online learning to adapt computation and data organization to available resources and emerging results
  - Transfer learning the knowledge of workflow optimization to other types of workflows and EH systems
  - **Ties to Managing System Resources**
- **Rapid adaptive analytics:** Data analytic workflows on EH systems must employ new, specialized algorithms and data organizations and facilitate use by domain scientists, not just data scientists.
  - Data and algorithm abstractions for domain scientists (e.g., probabilistic execution vs deterministic execution)
  - New data organizations and specialized robust algorithm developments for specific EH components
  - Important, but out of scope for this specific workshop

Current Status

- Draft report is under review, with revisions continuing
- PRD brochure has been drafted
- I am visiting Labs (BNL, ANL, and LBNL so far) and ASCR-supported facilities (ALCF, NERSC and ESnet so far), briefing about the workshop and inviting feedback



Thank you!

Lucy Nowell

lucy.nowell@science.doe.gov
https://science.energy.gov/ascr/

Basic Research Needs: Productive Computational Science in the Era of Extreme Heterogeneity

- *Reducing the time to scientific discovery in a transformational era of diverse scientific applications and revolutionary computer architectures.*
- *Revolutionizing how we utilize leadership class computing facilities for scientific innovation...*

Basic Research Needs: Productive Computational Science in the Era of Extreme Heterogeneity - 1/5

- **Maintaining and improving programmer productivity.**
- **Key question:** *What advances in programming models, environments, and tools are required to improve the productivity of a broad range of scientific software developers in the face of a increasingly diverse and complex computing environment?*

- The level of effort, time and expertise required to develop verifiable, high-performance scientific software is a key and often limiting factor in the overall time required to achieve key scientific milestones and transformational discoveries. Changes in supercomputer architecture greatly increase the difficulty and complexity of the process, therefore directly impacting the overall *time-to-solution*. Improved programmer productivity in terms of both reduced development time and increased application performance are key to maintaining competitive advantages.

Basic Research Needs: Productive Computational Science in the Era of Extreme Heterogeneity - 2/5

- **Managing resources intelligently.**
- **Key question:** *Can artificial intelligence and machine learning be effectively incorporated into system software to coordinate and control a large and diverse set of computing resources to meet the demands of increasingly complex scientific domains?*
- Manually coordinating and reasoning about data placement and the schedule of when operations occur quickly becomes intractable as machine complexity skyrockets. Leveraging the power of artificial intelligence and machine learning provide an opportunity for improving system utilization, overall productivity, and the pace of scientific discovery and innovation.

Basic Research Needs: Productive Computational Science in the Era of Extreme Heterogeneity - 3/5

- **Modeling & predicting performance.**
- **Key question:** *Can advanced modeling and simulation reveal the performance characteristic of emerging hardware technologies and predict how applications will perform on them, perhaps even guiding system designs to maximize the value of future systems for scientific discovery and innovation?*
- With a growing number of choices in types of processors, kinds of memories and storage architectures, the configuration of a high-performance supercomputer becomes increasingly challenging. The ability to evaluate potential system designs not only allows systems to be evaluated prior to procurement but also tailored and optimized to the needs of key applications, mission areas and use cases. This can not only improve overall productivity but reduce long-term system costs.
-

Basic Research Needs: Productive Computational Science in the Era of Extreme Heterogeneity - 4/5

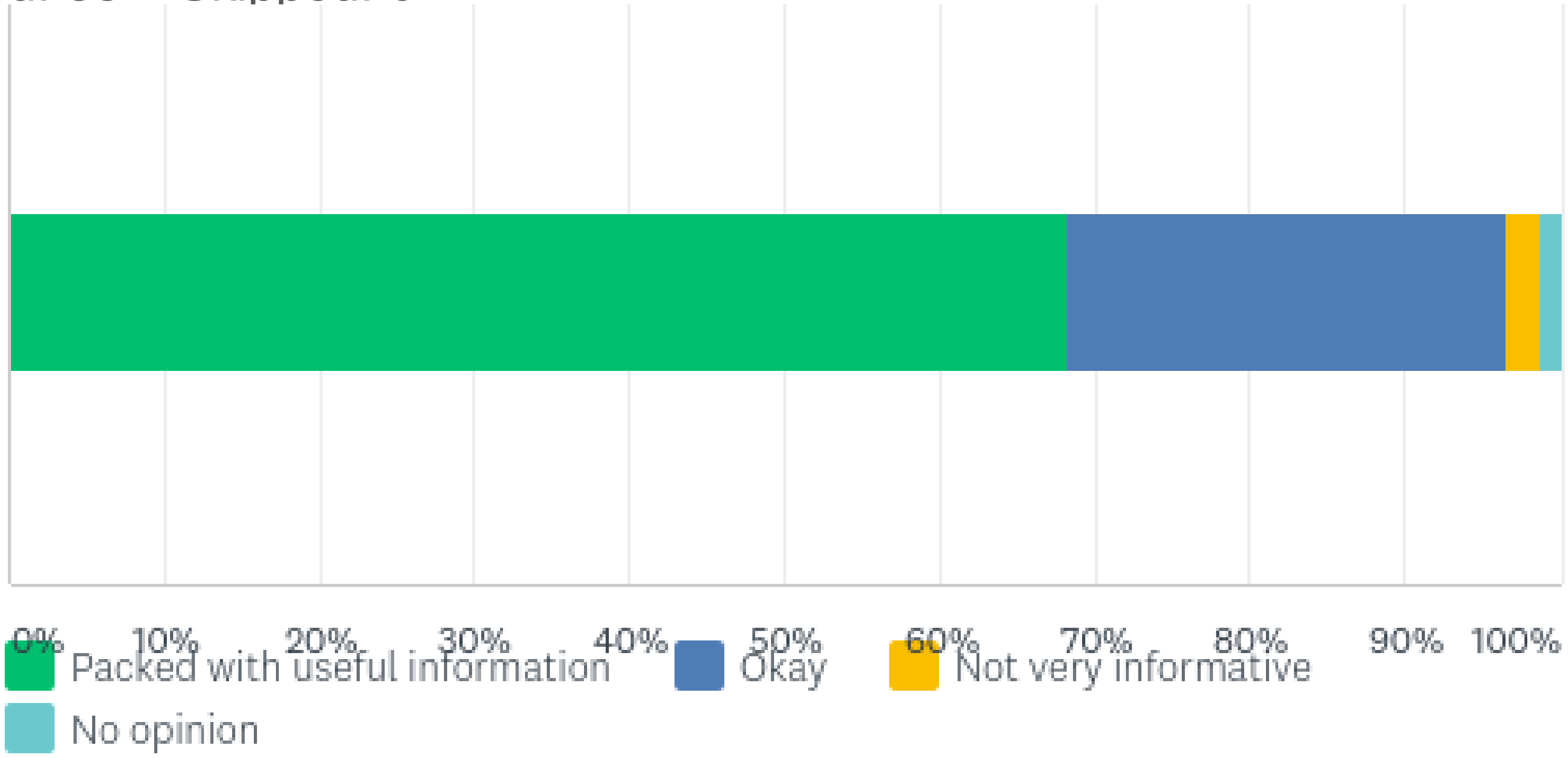
- **Enabling reproducible science despite diverse processors and non-determinism.**
- **Key question:** *What new methods and techniques will be required to support verifiable scientific findings on future systems that will be increasingly asynchronous and non-deterministic? What methods will be required to detect and mitigate pervasive faults and errors in an environment with increasingly diverse computing resources that can support trillion-way concurrency within acceptable power bounds?*
- The ability to validate scientific outcomes is essential, but it becomes exponentially more difficult when bit-wise reproducibility is no longer possible and hardware is increasingly unreliable. To maintain scientific accuracy and integrity, it is critical to be able to understand, reason about and evaluate the impact of these uncertainties in the execution of a given application and hardware combination.

Basic Research Needs: Productive Computational Science in the Era of Extreme Heterogeneity - 5/5

- **Facilitating Data Management, Analytics, and Workflows**
- **Key question:** *What system software infrastructure and tools will be necessary to achieve usable and productive scientific workflow in an ever increasingly complex computing environment?*
- A typical modern scientific study relies on multiple simulations and/or experiments, which must be coordinated across an increasing complex and often geographically distributed environment with engagement by diverse scientist. The selection of computing and data storage resources, cataloging of experimental results and findings, analysis of the resulting data, and the ability to track the process from formulation of a hypothesis to final discovery is already a difficult process. System software must facilitate the entire process, from resource discovery and scheduling to capturing the provenance that is essential to validate scientific findings.

Q1: Overall, I found the plenary talks to be

- Answered: 88   Skipped: 0



Q1: Overall, I found the plenary talks to be

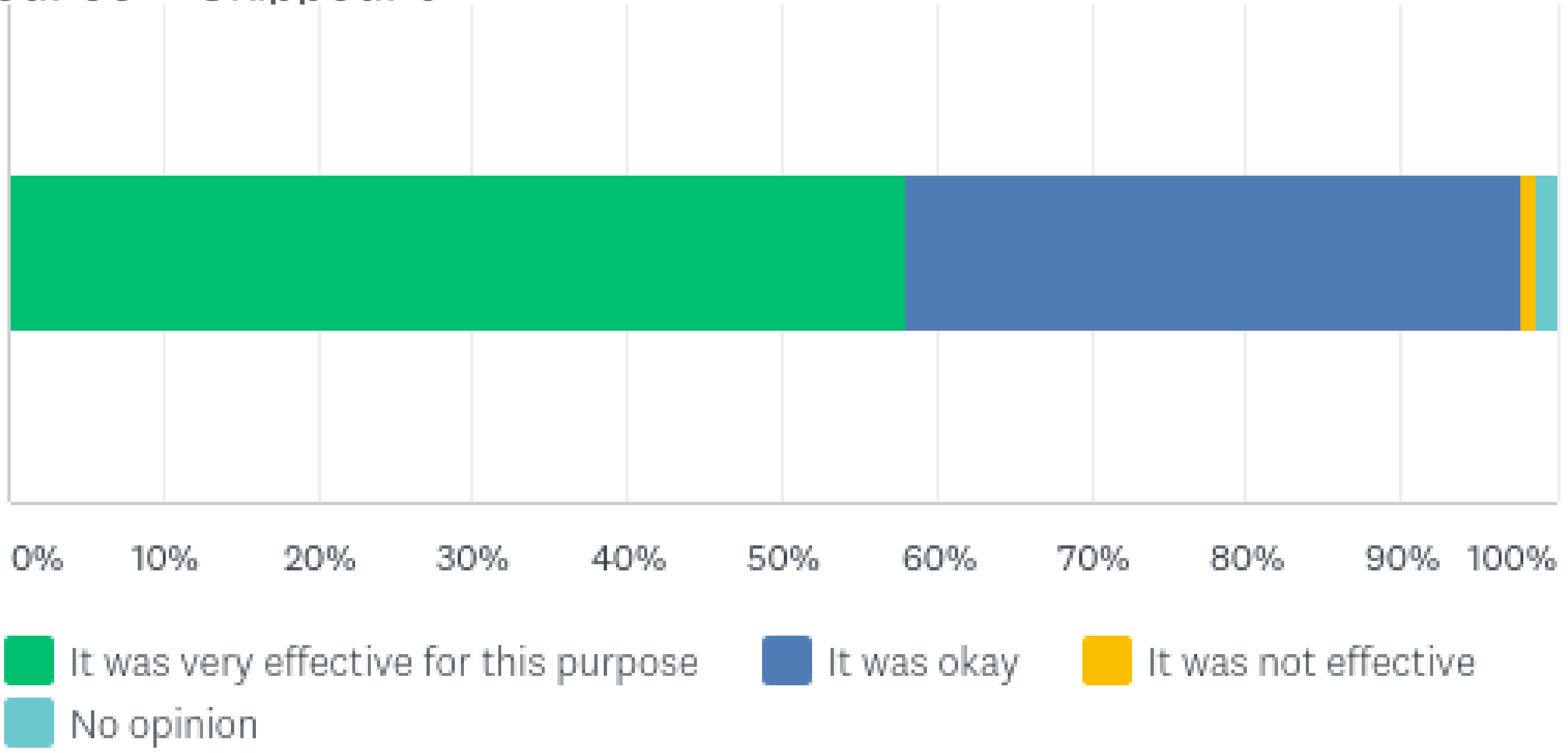
- Answered: 88   Skipped: 0

ANSWER CHOICES	RESPONSES	
Packed with useful information (1)	68.18%	60
Okay (2)	28.41%	25
Not very informative (3)	2.27%	2
No opinion (4)	1.14%	1
TOTAL		88

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	4.00	1.00	1.36	0.59

Q2: With regards to the Zoom technology used for the plenary sessions and some breakouts,

- Answered: 88   Skipped: 0



Q2: With regards to the Zoom technology used for the plenary sessions and some breakouts,

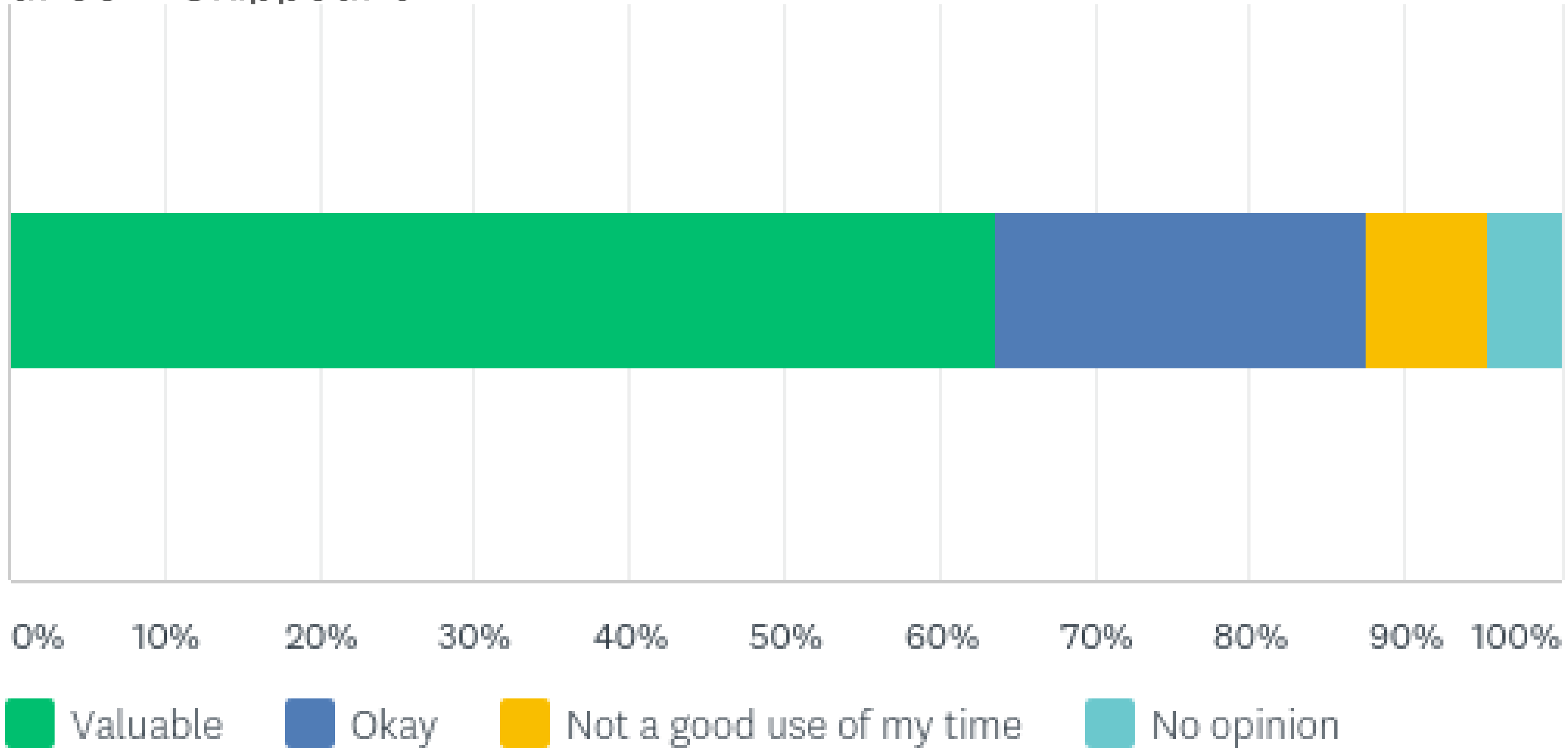
- Answered: 88 Skipped: 0

ANSWER CHOICES	RESPONSES	
It was very effective for this purpose (1)	57.95%	51
It was okay (2)	39.77%	35
It was not effective (3)	1.14%	1
No opinion (4)	1.14%	1
TOTAL		88

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	4.00	1.00	1.45	0.58

Q3: The breakout groups (BOGs) were

- Answered: 88   Skipped: 0



Q3: The breakout groups (BOGs) were

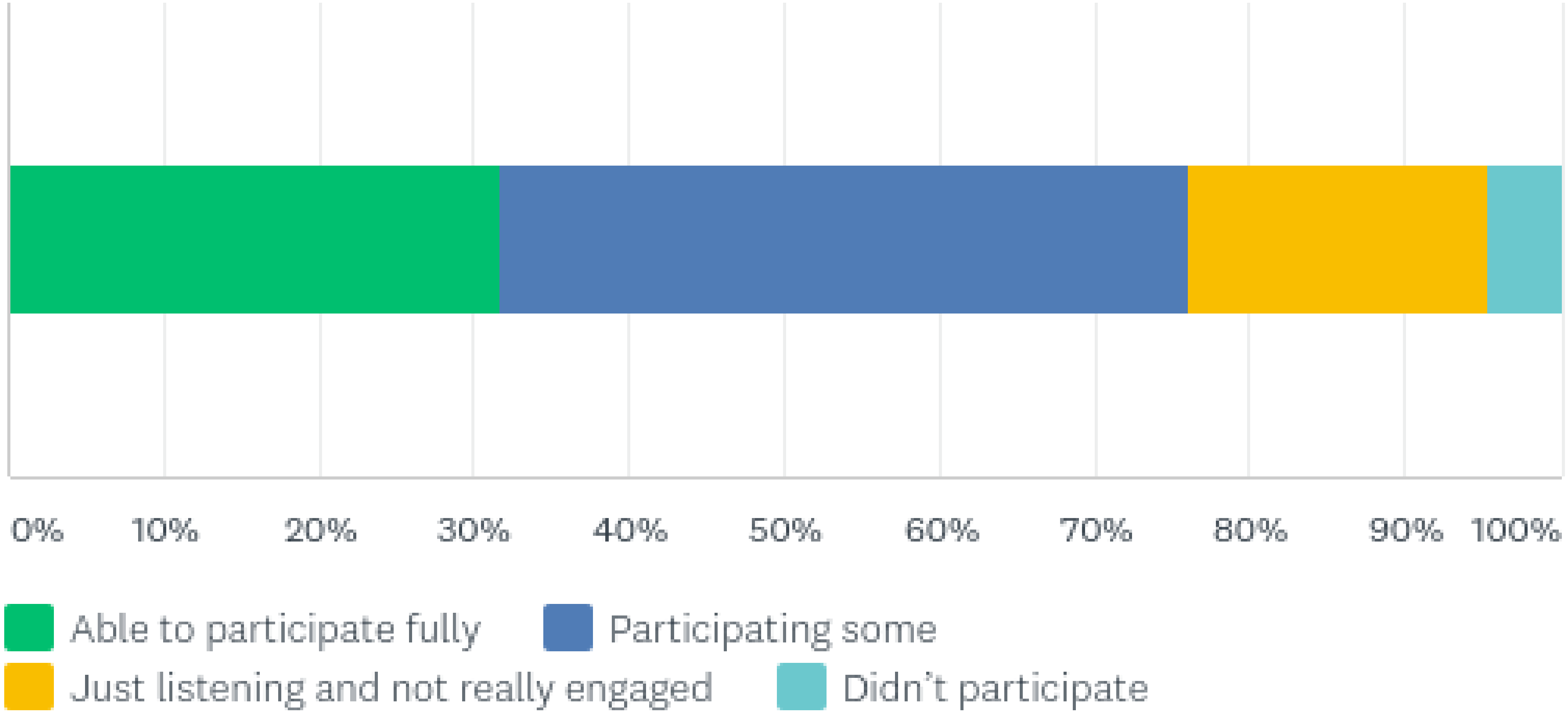
- Answered: 88   Skipped: 0

ANSWER CHOICES	RESPONSES	
Valuable (1)	63.64%	56
Okay (2)	23.86%	21
Not a good use of my time (3)	7.95%	7
No opinion (4)	4.55%	4
TOTAL		88

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	4.00	1.00	1.53	0.83

Q4: During breakout sessions (BOGs), I was

- Answered: 88   Skipped: 0



Q4: During breakout sessions (BOGs), I was

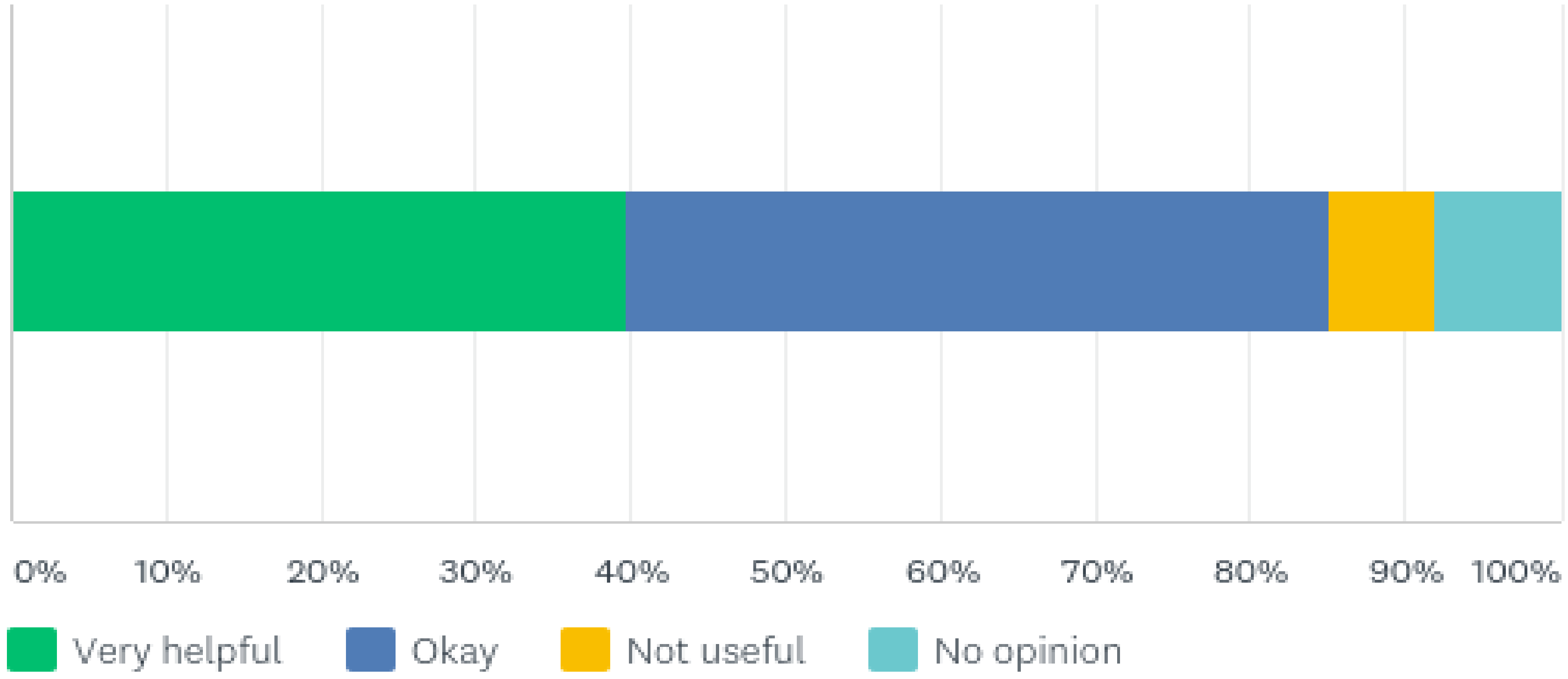
- Answered: 88   Skipped: 0

ANSWER CHOICES	RESPONSES	
Able to participate fully (1)	31.82%	28
Participating some (2)	44.32%	39
Just listening and not really engaged (3)	19.32%	17
Didn't participate (4)	4.55%	4
TOTAL		88

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	4.00	2.00	1.97	0.83

Q5: The report-back sessions were

- Answered: 88   Skipped: 0



Q5: The report-back sessions were

- Answered: 88   Skipped: 0

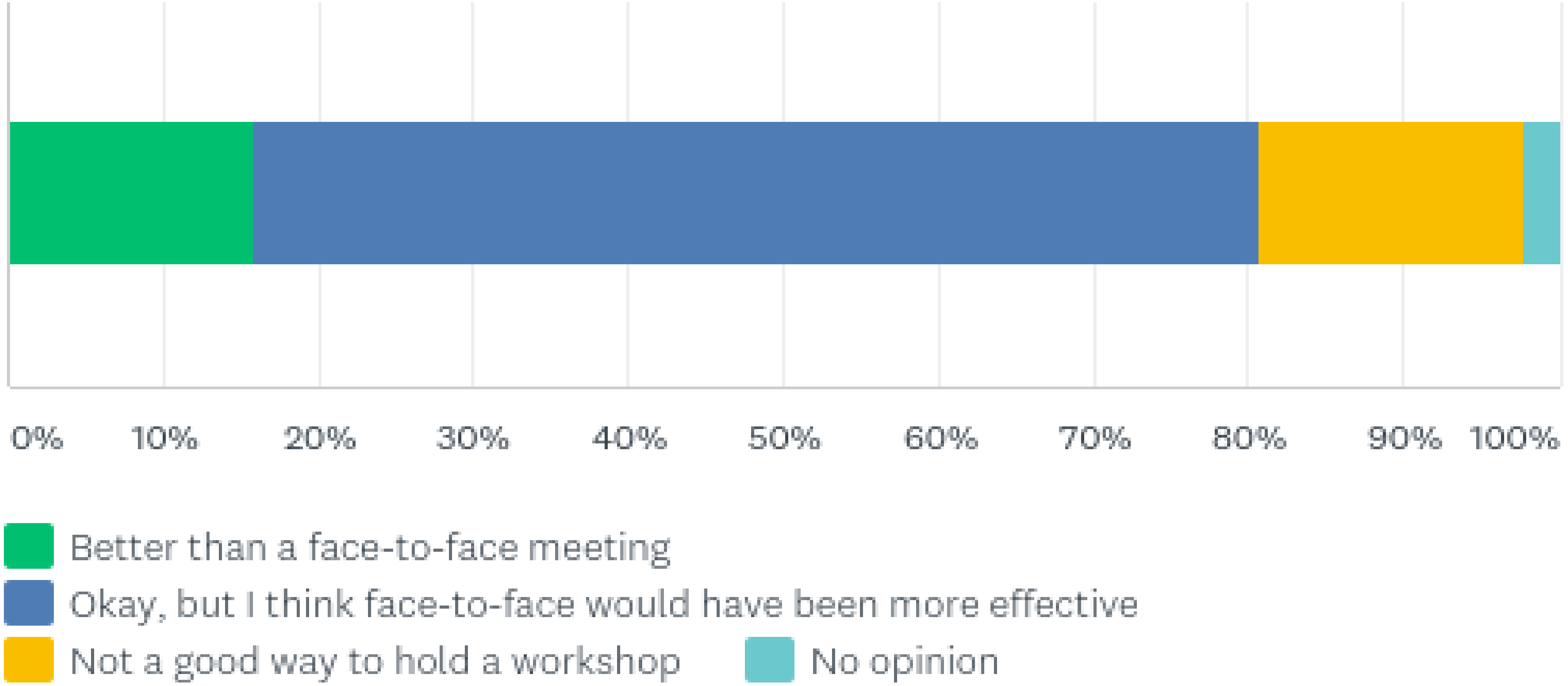
ANSWER CHOICES	RESPONSES	
Very helpful (1)	39.77%	35
Okay (2)	45.45%	40
Not useful (3)	6.82%	6
No opinion (4)	7.95%	7
TOTAL		88

BASIC STATISTICS

Minimum	Maximum	Median	Mean	Standard Deviation
1.00	4.00	2.00	1.83	0.87

Q6: I found the virtual meeting format to be

- Answered: 88   Skipped: 0



Q6: I found the virtual meeting format to be

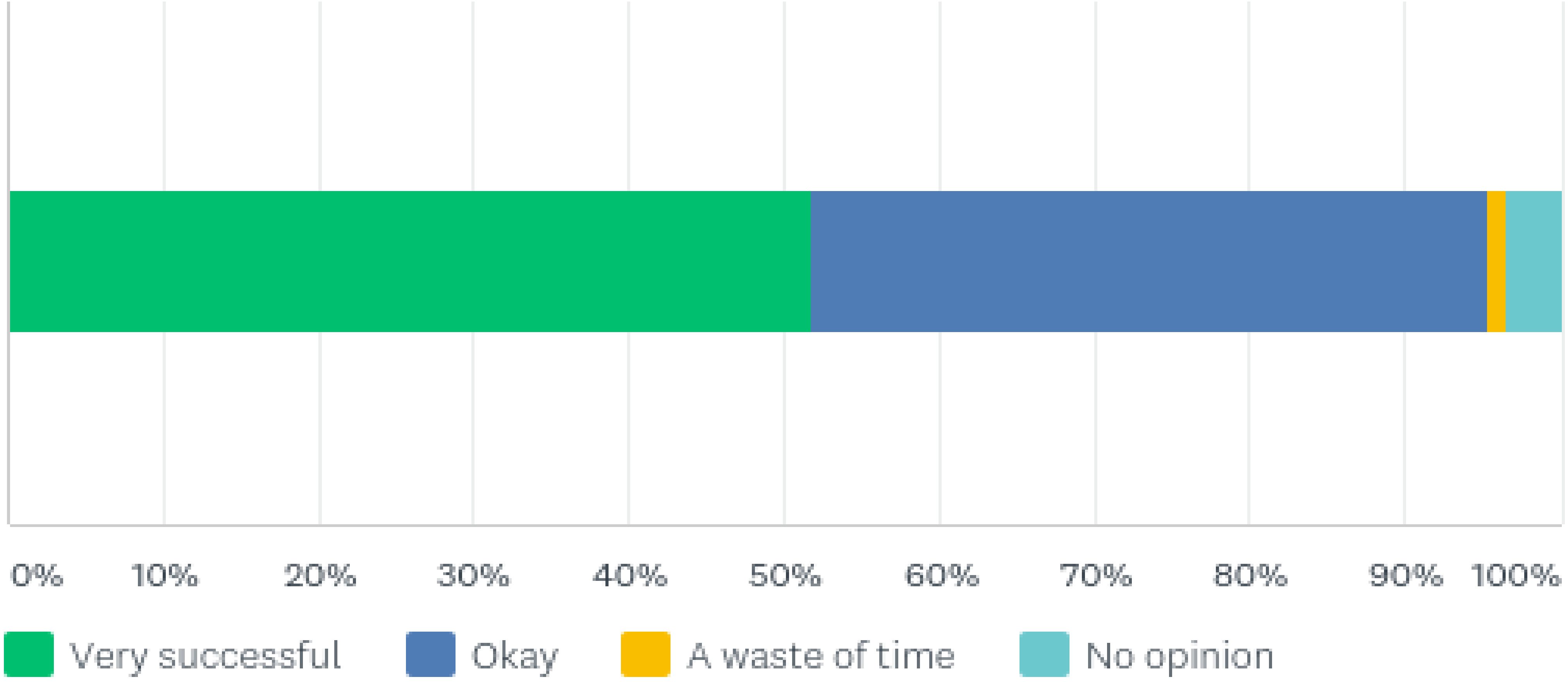
- Answered: 88   Skipped: 0

ANSWER CHOICES	RESPONSES	
Better than a face-to-face meeting (1)	15.91%	14
Okay, but I think face-to-face would have been more effective (2)	64.77%	57
Not a good way to hold a workshop (3)	17.05%	15
No opinion (4)	2.27%	2
TOTAL		88

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	4.00	2.00	2.06	0.65

Q7: Overall I found the Extreme Heterogeneity Workshop to be

- Answered: 87   Skipped: 1



Q7: Overall I found the Extreme Heterogeneity Workshop to be

- Answered: 87   Skipped: 1

ANSWER CHOICES	RESPONSES	
Very successful (1)	51.72%	45
Okay (2)	43.68%	38
A waste of time (3)	1.15%	1
No opinion (4)	3.45%	3
TOTAL		87

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	4.00	1.00	1.56	0.69