

Making Use of the Benefits of Domain Ignorance in RE

Ali Niknafs and Daniel M. Berry

David R. Cheriton School of Computer Science
University of Waterloo
Waterloo, Ontario, Canada

18 July 2013

First Observations of Benefits of Ignorance

Probably, the earliest observation of the benefits of ignorance was Burkinshaw's statement during the 1969 Second NATO Conference on Software Engineering:

*Get some **intelligent ignoramus** to read through your documentation and try the system; he will find many "holes" where essential information has been omitted. Unfortunately intelligent people don't stay ignorant too long, so ignorance becomes a rather precious resource. Suitable late entrants to the project are sometimes useful here.*

Our Previous Work

- Preliminary results were reported at the last RE Conference!
- We conducted a controlled experiment on 20 teams of all different combinations of DIs and DAs, testing

Main Hypothesis

A team consisting of a mix of DIs and DAs is more effective in an RE activity than is a team consisting of only DAs.

- We tested whether other factors, i.e., creativity, RE experience, and industrial experience, affected the results.

Our Previous Work

Main Conclusion w.r.t Main Hypothesis:

A team with a mix of domain familiarities *is* more effective in requirements idea generation than is a team with either

- only DIs or
- only DAs.

The Case Study

The case study reported in this RE conference was to corroborate the conclusions of the controlled experiment, by:

- getting one group with a mix of DAs and DIs to carry out the idea generation part of a requirements idea brainstorming session, and ...
- then asking the DA members of the group to compare the case study session with previous DA-only sessions.

Case Study Context

We carried out the idea generation part of a brainstorming session in a company, C , ...

- to generate requirement ideas for a system, S , ...
- situated in a specific domain, D , ...
- to be developed as one of C 's products.

Conclusions

- The DIs were generating out-of-the-box ideas.
- The DAs were interested in technical details, as they were seeking only implementable ideas. The DIs were oblivious to technical details and implementability.
- DAs are tied to solutions that they are already familiar with. One of the DAs said some of the ideas that DIs brought up were irrelevant to the problem. However, the DAs were often able to make an irrelevant idea worthy of follow up by modifying the idea to fit what they perceive as the domain D of C .

Conclusions

- Considering that the main purpose of the idea generation part of brainstorming is to generate as many ideas as possible, what the DA observed is about normal.
- There were indications that the DIs may have generated some ideas that were innovative to *C*. (Because of confidentiality they could not tell us for sure.)
- Ideas can be borrowed from domains different from *D*.
- Finally, the experience suggest that in company *C*, brainstorming groups should be composed of domain experts and new employees.

Expected Application of the Results

Help RE managers in forming teams that are performing knowledge-intensive RE activities, by

- providing a list of RE activities for which domain ignorance is at least helpful and
- providing advice on the best mix of DIs and DAs for any RE activity.
- providing a useful role for new hires that allows them to be productive from the start while learning about the domain slowly without being a time drain on their mentors.

Expected Application of the Results

Make a niche business of supplying intelligent ignoramuses to participate in clients' RE teams.

Hypothesis

Main Hypothesis

A team consisting of a mix of DIs and DAs is more effective in an RE activity than is a team consisting of only DAs.

Lessons Learned from Pilot Studies

- 1 Find a suitable problem domain.
- 2 Consider other factors (e.g. industrial experience) in analyzing the results.
- 3 Assess also the quality of the DGUs.
- 4 For many domains, so-called DIs turn out not to be real DIs, and so-called DAs turn out not to be real DAs.

Lessons Learned from Pilot Studies

Lessons 1 and 4 taught us that we need a problem domain that *partitions* the set of subjects with precision into

- DAs
- DIs

with no one in between.

We thought very hard to find such a domain, bidirectional word processing:

- CSers from the Middle East are DAs.
- CSers from elsewhere are DIs.

Experiment Context

- *GT*: The first, idea-generation step in a brainstorming activity to generate requirement ideas for a CBS.
- *DGUs*: Requirement ideas
- *Domain*: Bidirectional word processing
- *Subjects*: Volunteer subjects were recruited from a “Software Requirements and Specification” course and from outside the course, but nevertheless in CS or a related discipline.
- *Teams*:
 - 3I: a team consisting of 3 DIs and 0 DAs,
 - 2I: a team consisting of 2 DIs and 1 DAs,
 - 1I: a team consisting of 1 DI and 2 DAs,
 - 0I: a team consisting of 0 DIs and 3 DAs.

Variables

- Independent Variables about a team
 - *Mix of Domain Familiarities*
 - *Creativity Level*
 - *RE Experience*
 - *Industrial Experience*
- Dependent Variable
 - *Effectiveness*

Evaluation of Generated Ideas

- The quantitative data is the number of raw ideas generated by each team, which is a good measure for the GT = brainstorming (because quantity is the *goal* of the first stage of brainstorming).
- To better compare the performance of the teams, Niknafs considered also the quality of their generated ideas.

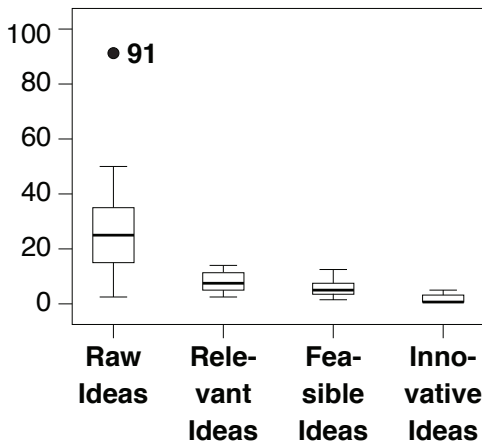
Quality of Generated Ideas

Each idea is classified according to three characteristics:

- 1 *Relevancy*: an idea is considered relevant if it has something to do with the domain.
- 2 *Feasibility*: an idea is considered feasible if it is relevant and it is correct, well presented, and implementable.
- 3 *Innovation*: an idea is considered innovative if it is feasible and it is not already implemented in an existing application for the domain known to the evaluator.

Outliers

- Boxplots were used to graphically expose any outliers.



ANOVA Prerequisites

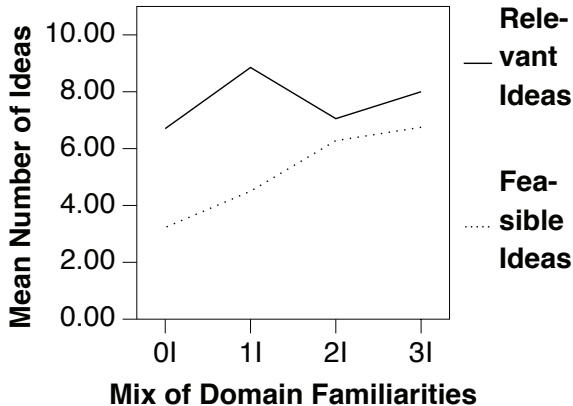
- An ANOVA was applied to the dependent variables whose values met the prerequisites for an ANOVA; i.e. the numbers of generated raw, relevant, and feasible ideas.
- For innovative ideas, another, non-parametric test was used.

ANOVA Results

Effect	Raw Ideas				Relevant Ideas				Feasible Ideas			
	<i>F</i>	<i>p</i>	<i>f</i> ²	<i>P</i>	<i>F</i>	<i>p</i>	<i>f</i> ²	<i>P</i>	<i>F</i>	<i>p</i>	<i>f</i> ²	<i>P</i>
Mix of Domain Familiarities	.165	.915	.011	.068	8.675	.032	.319	.816	13.486	.015	.449	.941
Creativity	.921	.469	.048	.146	3.918	.114	.159	.459	.984	.449	.051	.153
Industrial Experience	.563	.609	.031	.107	10.089	.027	.331	.833	4.381	.098	.173	.499
RE Experience	.145	.722	.008	.063	.173	.699	.009	.65	.035	.861	.002	.53

F is *F*-test; *p* is *p*-value of *F*-test; *f*² is Cohen effect size; *P* is post-hoc power.

ANOVA Results: Impact of Domain Knowledge



Threats to Validity

- *Conclusion Validity*: Low Statistical Power: 20 teams would be enough to achieve statistical power of 0.80, but, the unequal number of teams in the mixes reduces statistical power.
- *Internal Validity*: Voluntary Subjects: All subjects were voluntary but were randomized to the extent possible while still getting the necessary mixes of domain familiarities among the teams.

Threats to Validity

- *Construct Validity*: Confounding Constructs: Sometimes the value of an independent variable affects the results more than the presence or absence of the variable would.
- *External Validity*: Population Validity: The experiment used student subjects instead of professional analysts, although the students are mostly co-op and work one term per year.

Conclusion About Hypotheses

- The main hypothesis is accepted:
The effectiveness of a team in requirements idea generation **is** affected by the team's mix of domain familiarities.

Main Result

From these results, considering the threats, the main hypothesis, that

A team consisting of mix of DIs and DAs is more effective in requirements idea generation than a team consisting of only DAs,

appears to be weakly supported.