Validation of the Effectiveness of an Optimized EPMcreate as an Aid for Creative Requirements Elicitation

Victoria Sakhnini[1],

Daniel M. Berry[1], & Luisa Mich[2]

[1]University of Waterloo, Canada
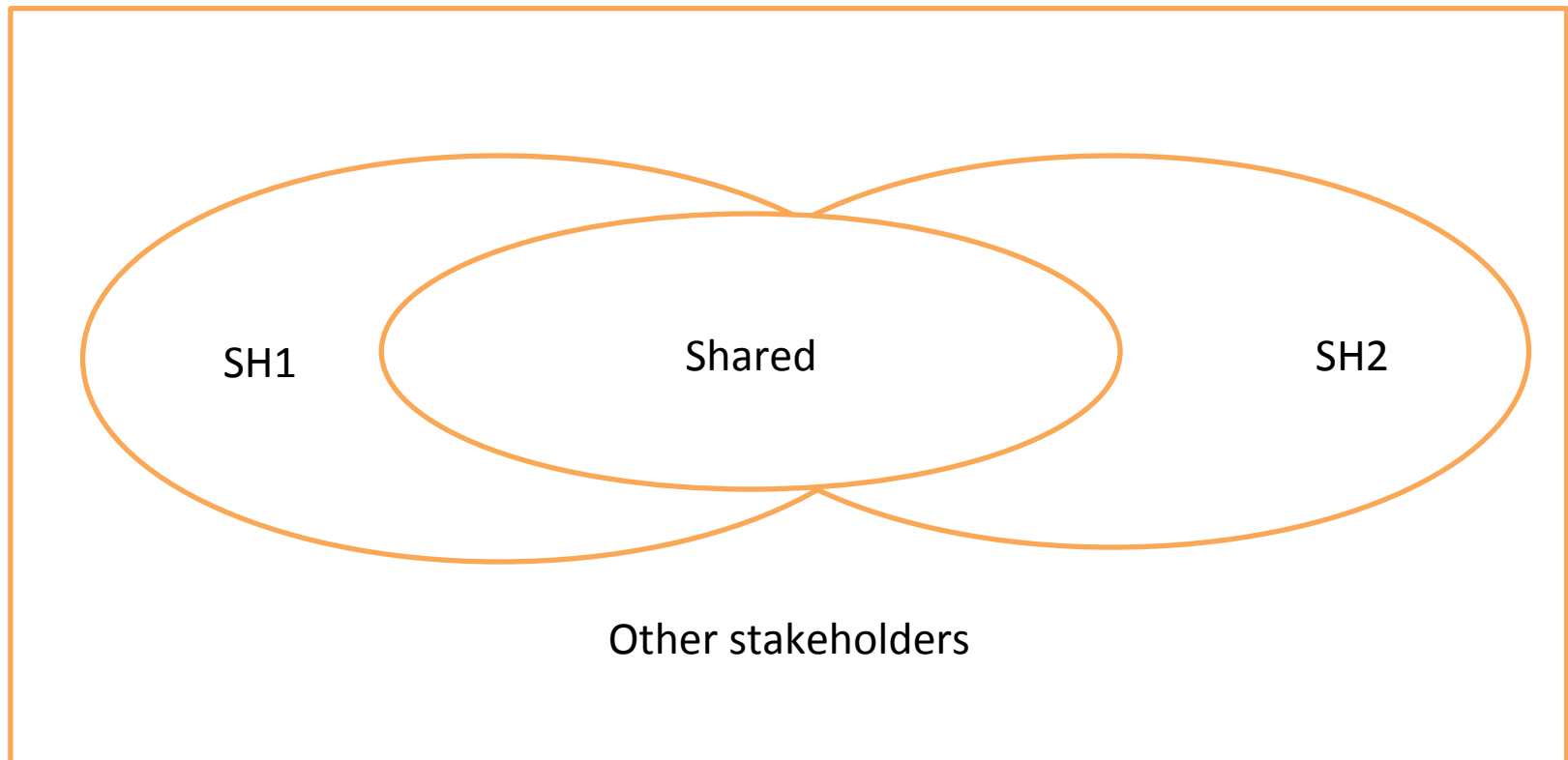vsakhnin@uwaterloo.ca, dberry@uwaterloo.ca
[2] University of Trento, Italy
luisa.mich@unitn.it

The EPMcreate Technique

- EPMcreate supports idea generation by focusing the search for ideas on only one logical combination of two stakeholders' viewpoints at a time.

- Sixteen such combinations are possible …

The EPMcreate Technique (cont.)

The EPMcreate Technique (cont.)

- each combination corresponding to one of the Boolean functions, fi for $0 \leq i \leq 15$, of two variables.

f0 represents disagreeing with everything, regardless of what SH1 and SH2 want

V1	V2	C
0	0	0
0	1	0
1	0	0
1	1	0

f0 = nothing

f1 represents when what SH1 wants and what SH2 wants coincide.

V1	V2	C
0	0	0
0	1	0
1	0	0
1	1	1

$$f1 = SH1 \wedge SH2$$

f2 represents when what SH1 wants and what SH2 does not want coincide.

V1	V2	C
0	0	0
0	1	0
1	0	1
1	1	0

$$f2 = SH1 \wedge \neg SH2$$

f3 represents what SH1 wants, regardless of what SH2 wants.

V1	V2	C
0	0	0
0	1	0
1	0	1
1	1	1

f3 = SH1

f5 represents what SH2 wants, regardless of what SH1 wants

V1	V2	C
0	0	0
0	1	1
1	0	0
1	1	1

f5 = SH2

f10 represents what SH2 does not want, regardless of what SH1 wants.

V1	V2	C
0	0	1
0	1	0
1	0	1
1	1	0

$$f10 = \neg\ SH2$$

Power Only EPMcreate

The optimization of EPMcreate is called the "Power-Only EPMcreate (POEPMcreate)", because it does only the four steps whose names are powers of two, namely f1, f2, f4, and f8.

Power Only EPMcreate (cont.)

f1		
V1	V2	C
0	0	0
0	1	0
1	0	0
1	1	1

f2		
V1	V2	C
0	0	0
0	1	0
1	0	1
1	1	0

f4		
V1	V2	C
0	0	0
0	1	1
1	0	0
1	1	0

f8		
V1	V2	C
0	0	1
0	1	0
1	0	0
1	1	0

f1= SH1 ^ SH2   f2= SH1 ^ ¬SH2

f4= ¬SH1 ^ SH2   f8= ¬SH1 ^ ¬SH2

$$f1 = SH1 \wedge SH2 \quad f2 = SH1 \wedge \neg SH2$$

$$f4 = \neg SH1 \wedge SH2 \quad f8 = \neg SH1 \wedge \neg SH2$$

SH1

Shared

SH2

Other stakeholders
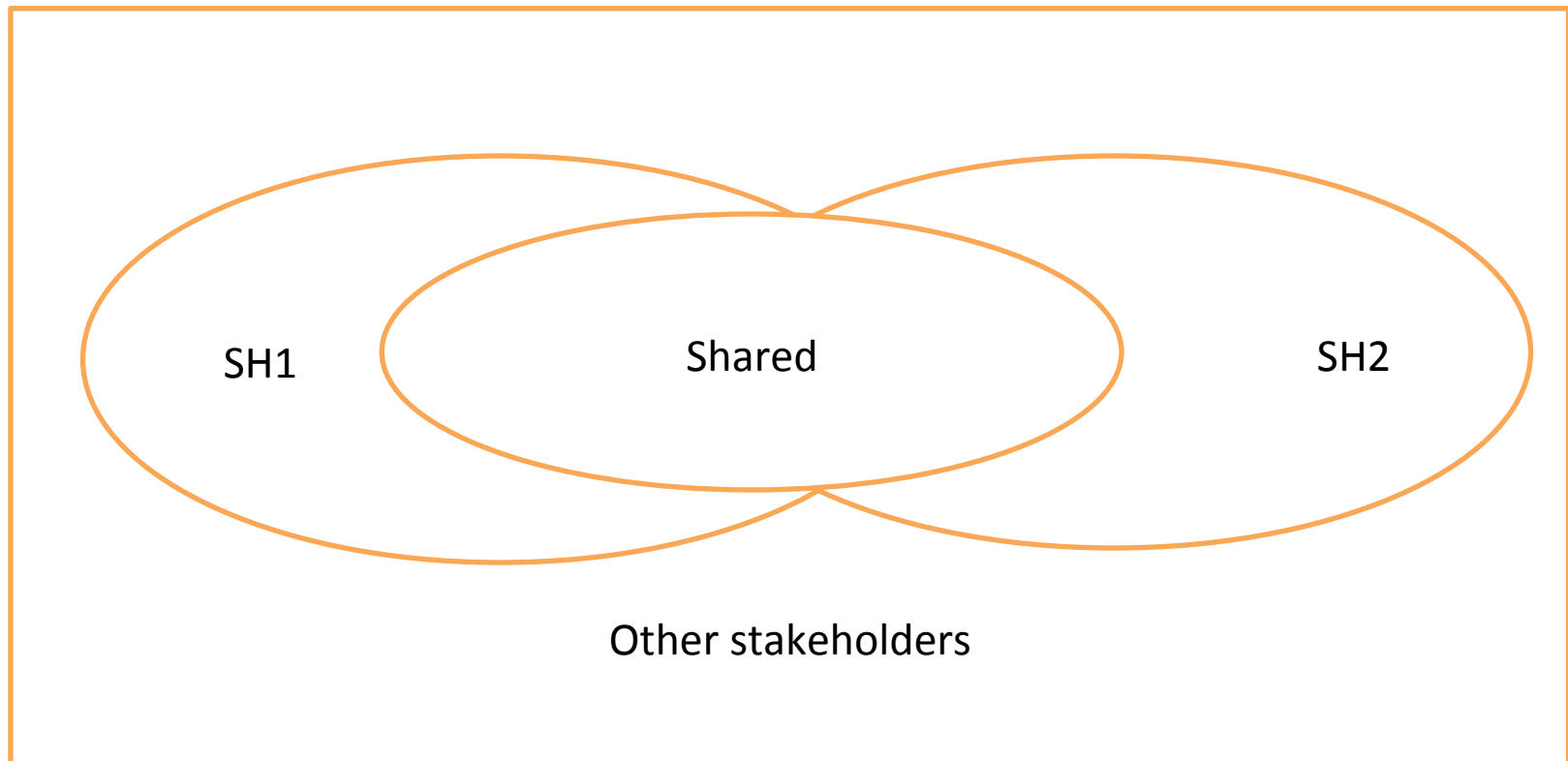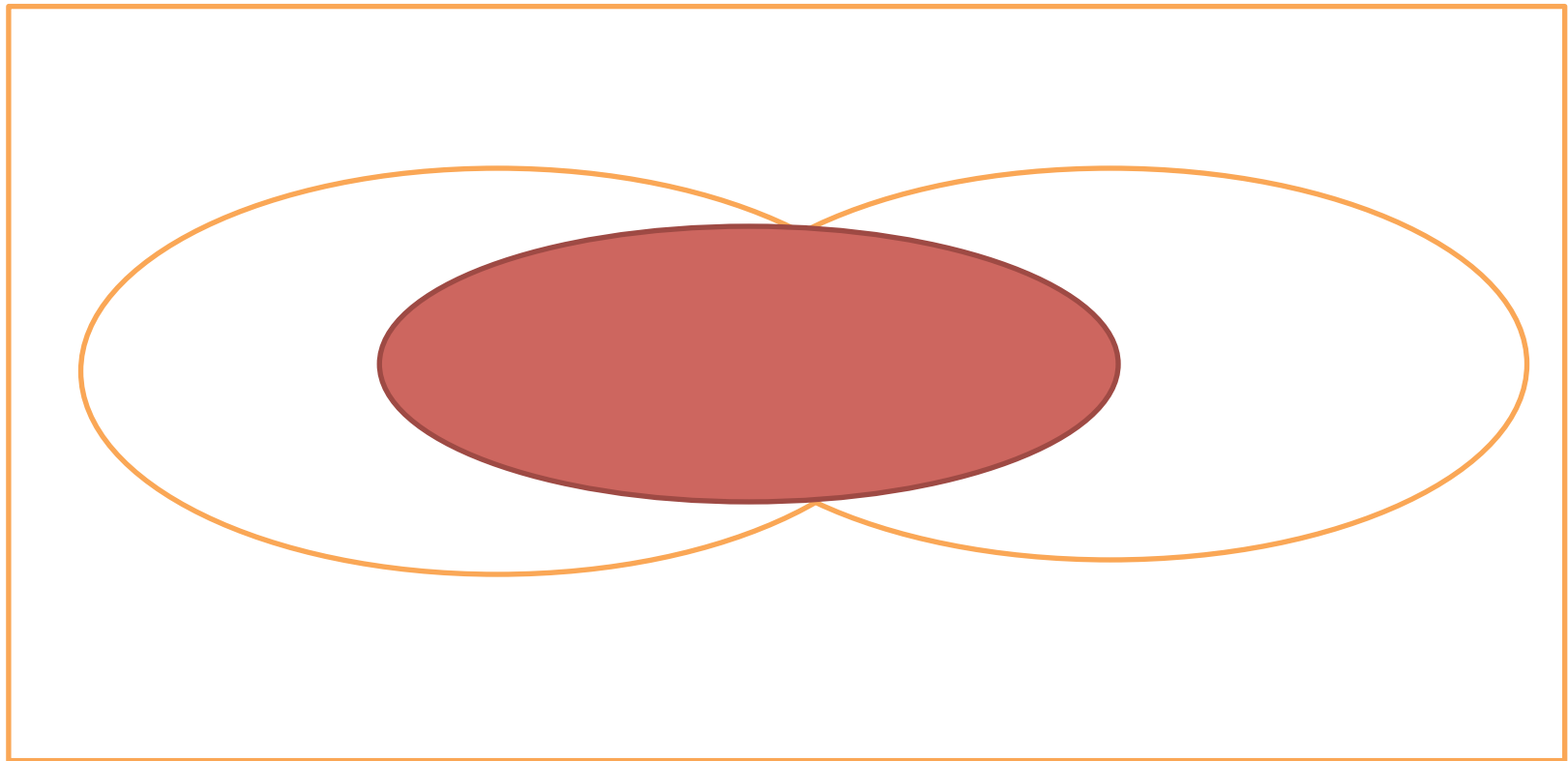
f1= SH1 ^ SH2   f2= SH1 ^ ¬SH2
f4= ¬SH1 ^ SH2   f8= ¬SH1 ^ ¬SH2
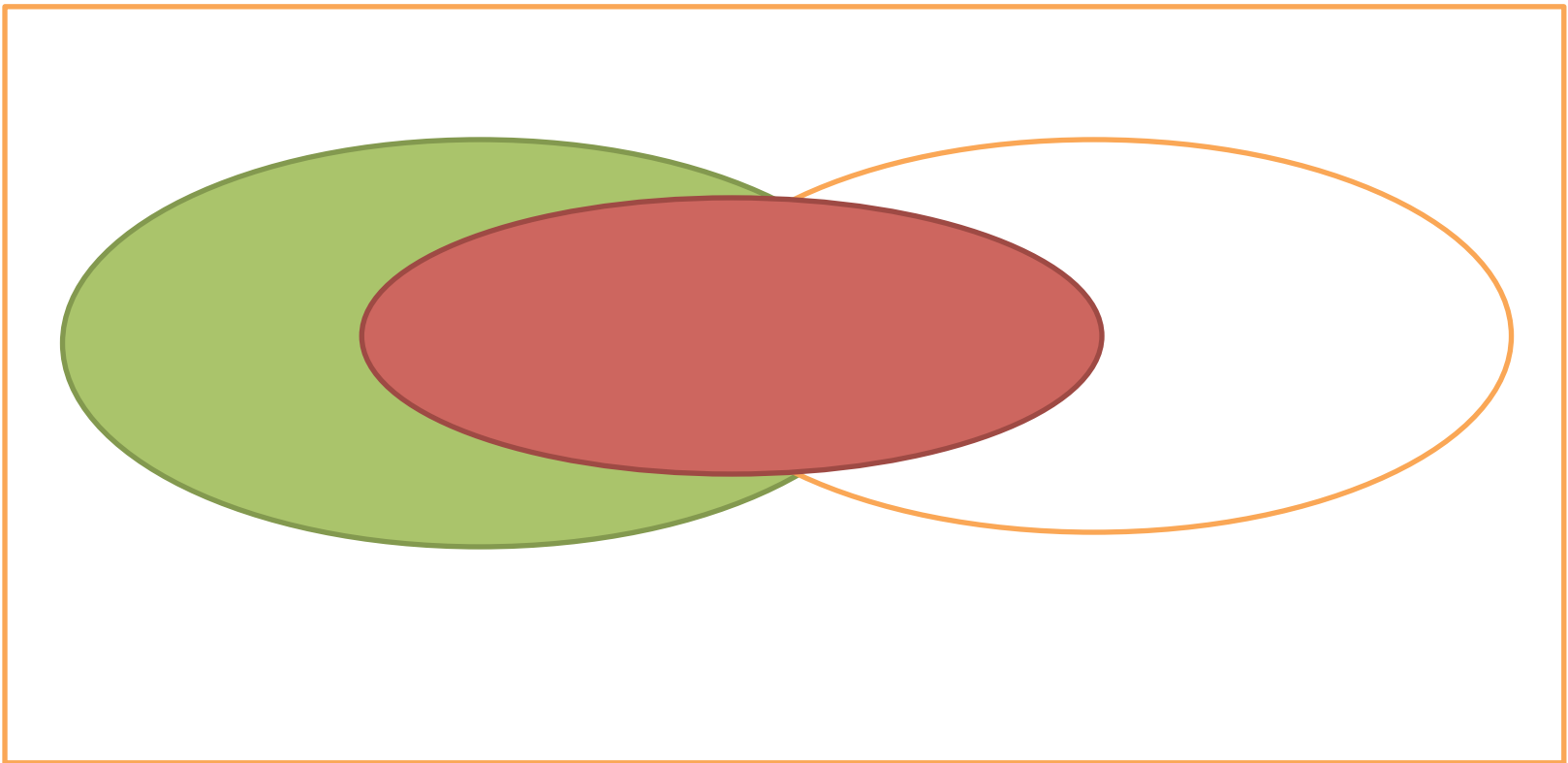
f1= SH1 ^ SH2   f2= SH1 ^ ¬SH2
f4= ¬SH1 ^ SH2   f8= ¬SH1 ^ ¬SH2

f1= SH1 ^ SH2   f2= SH1 ^ ¬SH2
f4= ¬SH1 ^ SH2   f8= ¬SH1 ^ ¬SH2

f1= SH1 ^ SH2   f2= SH1 ^ ¬SH2
f4= ¬SH1 ^ SH2   f8= ¬SH1 ^ ¬SH2

Main objective

To demonstrate the effectiveness of POEPMcreate as a creativity enhancement technique (CET).

Experiment Design and Planning

- We conducted an experiment and compared the requirement ideas for one CBS generated by six groups, two of which used POEPMcreate, two of which used EPMcreate, and two of which used brainstorming.

- The same number of subjects participated in the experiment for the same amount of time in each group.

Experiment Design and Planning (cont.)

- Each group was to generate, using its assigned CET, ideas for requirements for an improved version of an existing Web site.

- The Web site was that of a Canadian high school with information directed to students, parents, teachers, and administrators.

Rationale for Choosing This Site

The site was chosen for

1. its accessibility,

2. lack of intellectual property restrictions, and

3. the fact that as educators, the authors could be considered domain experts.

Stakeholders for EPMcreate and POEPMcreate

The two types of stakeholders whose viewpoints would be adopted by the EPMcreate and POEPMcreate groups were

students

and

parents.

Hypotheses

H1: The POEPMcreate is more effective than the full 16-step EPMcreate in helping to generate requirement ideas.

H2: The full 16-step EPMcreate is more effective than brainstorming in helping to generate requirement ideas.

Measuring the Effectiveness of a Creativity Enhancement Technique (CET)

The effectiveness of a CET is measured by two numbers about the ideas generated when using the CET,

1. the quantity, i.e., the raw number, of ideas and

2. the number of high quality ideas.

Quality of an Idea

1. new and realizable ⬅ Highest

2. new and not realizable

3. not new and not realizable

4. not new and realizable ⬅ Lowest

To evaluate the quality of the ideas:

- Two domain experts independently classified each idea into one of 4 rankings.

- But the authors could be biased towards acceptance of the hypotheses!

Avoiding Bias

- Two domain experts independently classified each idea into one of 4 rankings.

- We merged the requirement ideas generated by the 6 groups into one file.

- We then sorted the ideas alphabetically to produce the list of ideas to be evaluated.

Avoiding Bias (cont.)

- Each evaluator assigned a ranking to each idea.

- The rankings were copied to the original idea files, in order to be able to evaluate the quality of the requirement ideas of each group separately.

Steps of the Experiment

Step 1: 20 minutes for each subject to filling a general information form, to allow us to know his or her background:

Step 1 (cont.)

Step 1: … The form included questions about his or her

age, gender, native language,

computer science (CS) courses, qualifications related to CS, employment history in CS, and

knowledge of the CETs: brainstorming, EPMcreate, and POEPMcreate.

Steps of the Experiment (cont.)

Step 2: 30 minutes for each subject to take a creativity assessment test to measure his or her native creativity.

Step 3: 10 minutes for us to deliver to each group an explanation about the experiment and the CET it was to use.

Steps of the Experiment (cont.)

Step 4: 120 minutes for each group to carry out its requirements elicitation session using the group's CET.

Assigning Subjects into Balanced Groups

Group	Technique	# Males	# Fe-males	# native in Eng-lish	# not native in Eng-lish	# taken ≥ 10 CS courses	# taken 3–5 CS courses	# worked profes-sion-ally	# not worked profes-sion-ally	Aver-age age	Aver-age Wil-liams test score
1	POEPMcreate	3	1	1	3	2	2	2	2	25.5	70.66
2	POEPMcreate	2	2	2	2	2	2	3	1	23.8	71.00
3	EPMcreate	2	2	2	2	1	3	3	1	21.5	70.75
4	EPMcreate	3	1	1	3	2	2	1	3	23.4	70.60
5	Brainstorming	4	0	3	1	1	3	1	3	20.2	71.60
6	Brainstorming	2	2	2	2	3	1	3	1	25	70.25

Evaluation of the Quantity of the Requirement Ideas

Graph of the number of ideas generated
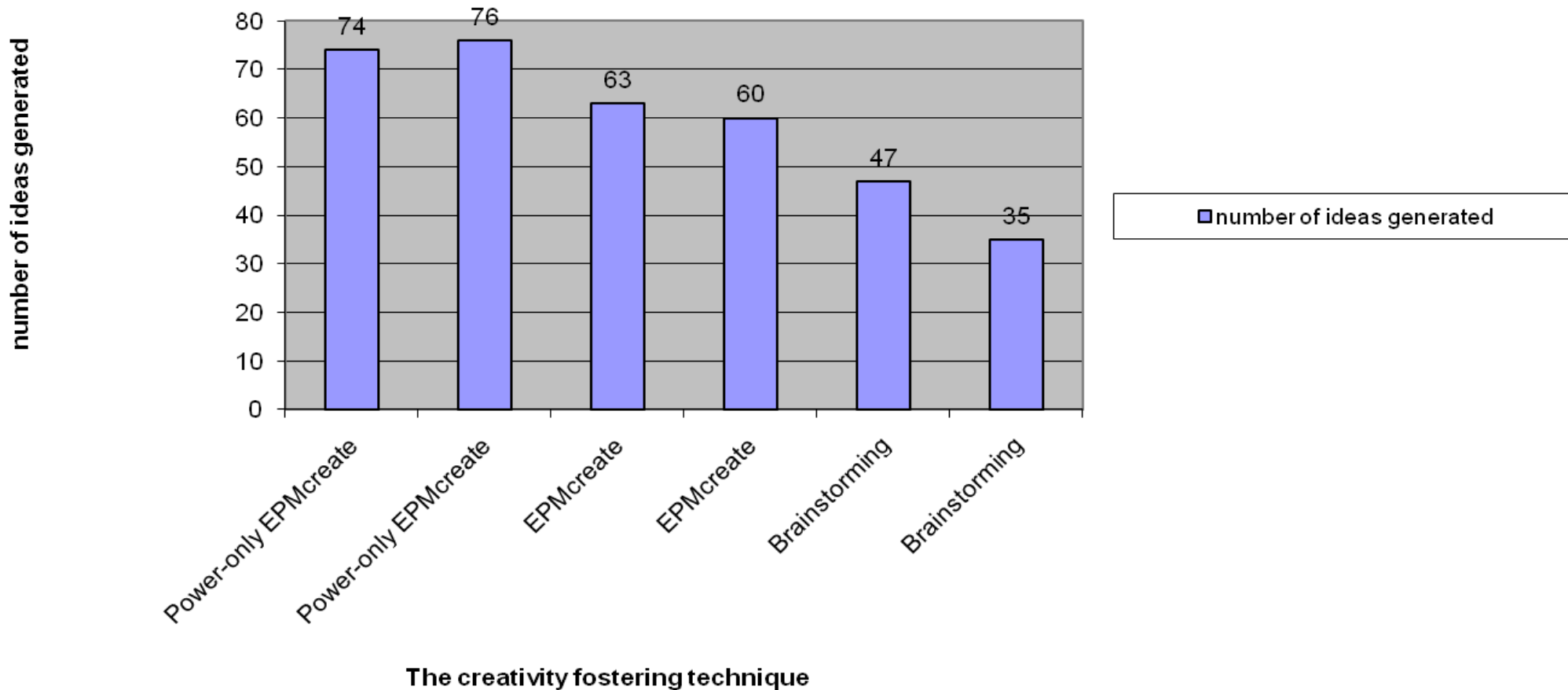
A two-sample T-test for unequal variances

POEPMcreate   >   EPMcreate

helps generate a larger

Quantity of requirement

ideas than

$\alpha = 0.05$ P $= 0.0087$

A two-sample T-test for unequal variances

POEPMcreate   >   EPMcreate

α = 0.05   P = 0.0087

EPMcreate   >   brainstorming

α = 0.10   P = 0.088.

A two-sample T-test for unequal variances

POEPMcreate   >   EPMcreate

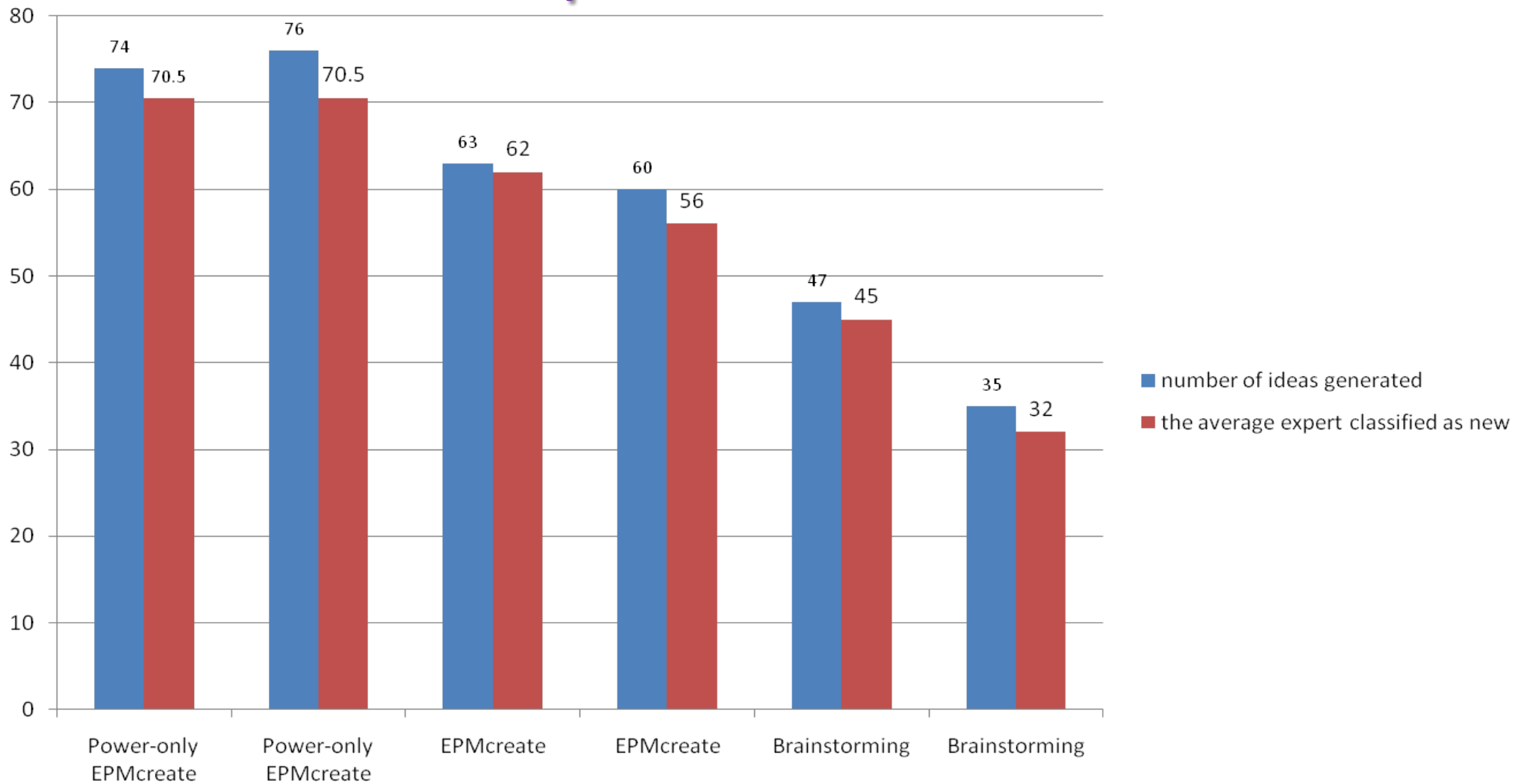$\alpha = 0.05$ P = 0.0087

EPMcreate   >   brainstorming

$\alpha = 0.10$ P = 0.088.

POEPMcreate   >   brainstorming

$\alpha = 0.06$ P = 0.053.

Evaluation of the Quality of the Requirement Ideas

A two-sample T-test for unequal variances

POEPMcreate   >   EPMcreate

helps generate more

high quality

requirement ideas than

α = 0.085 P = 0.081.

A two-sample T-test for unequal variances

POEPMcreate   >   EPMcreate

α = 0.085 P = 0.081.

EPMcreate   >   brainstorming

α = 0.12 P = 0.11.

A two-sample T-test for unequal variances

POEPMcreate   >   EPMcreate

α = 0.085 P = 0.081.

EPMcreate   >   brainstorming

α = 0.12 P = 0.11.

POEPMcreate   >   brainstorming

α = 0.07 P = 0.064.

Analysis of Corroboratory Data

- More than one POEPMcreate user said that the POEMcreate is easy to apply and the four foci were helpful.

- One EPMcreate user complained of having to rush and to change focus 16 times.

Analysis of Corroboratory Data (cont.)

- More than one POEPMcreate and EPMcreate user said that they felt productive and that they felt that they had not missed anything as they might have in brainstorming.

- More than one brainstorming user said that they felt unfocused, that they jumped among ideas, and that they might have missed ideas.

Threats to Validity and Limitations

- In spite of statistical significance, The small number of data points stands in the way of whether the results can be generalized to other cases, with different kinds of subjects, with different kinds of CBS.

- Small number of points also increases chances of a false positive result.

- Would other choices of stakeholders, e.g., of teachers, work as well?

Future work

- The results are strong enough that it is worth conducting more experiments to test these hypotheses, with more subjects and different CBSs about which to generate requirement ideas.

- To evaluate the effectiveness of other optimizations of EPMcreate and of other orderings of the steps of EPMcreate, POEPMcreate, and the other optimizations.

Future work (cont.)

- To compare the effectiveness of POEPMcreate applied by individuals to the effectiveness of POEPMcreate applied by groups.
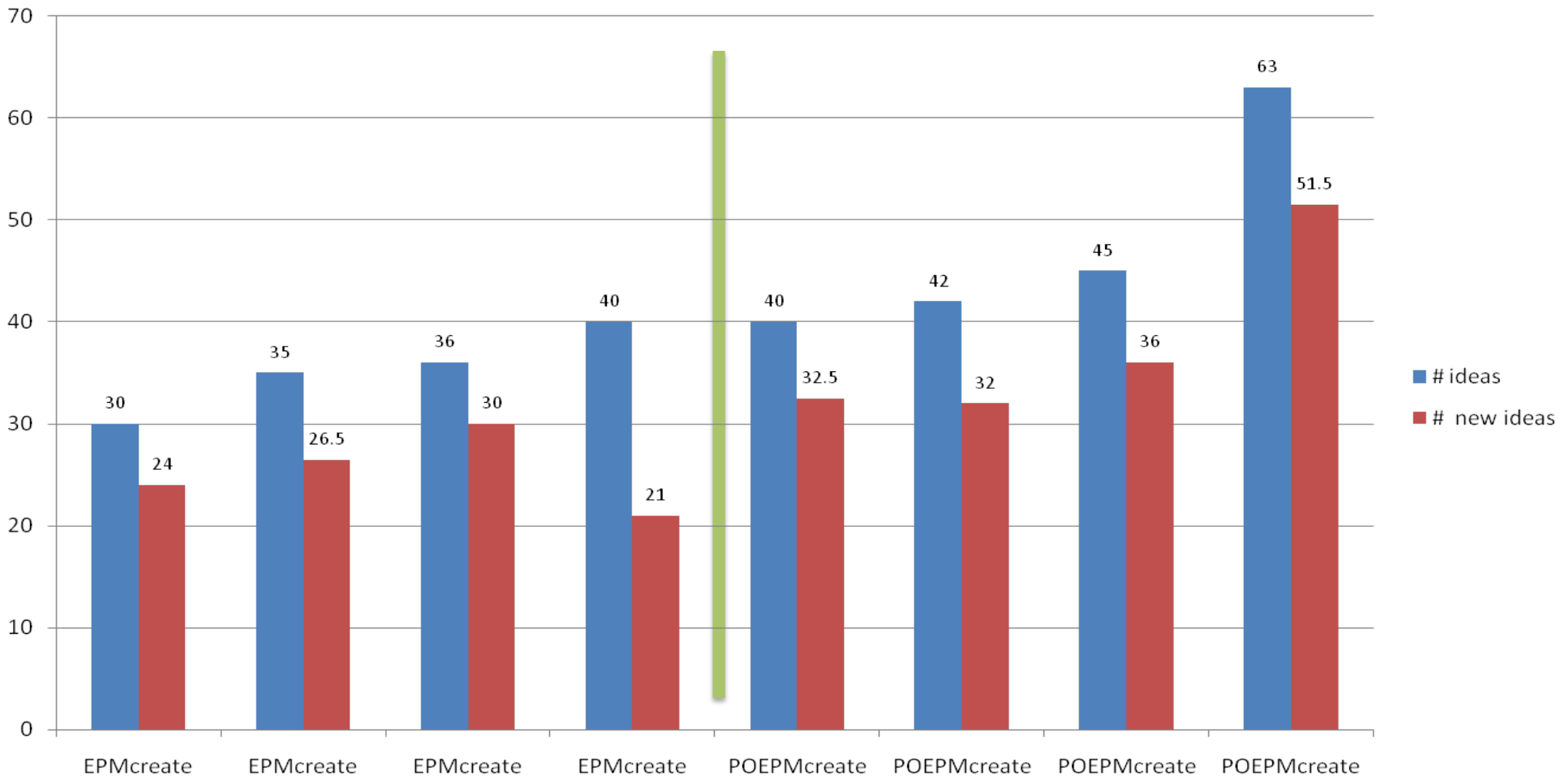
Follow up study

Since preparing this paper, we have done a follow up study to address the threats arising from having too few data points.

We did another experiment with same Web site, just EPMcreate and POEPMcreate, 4 teams for each method, but only 2 members per team.

Follow up study (cont.)
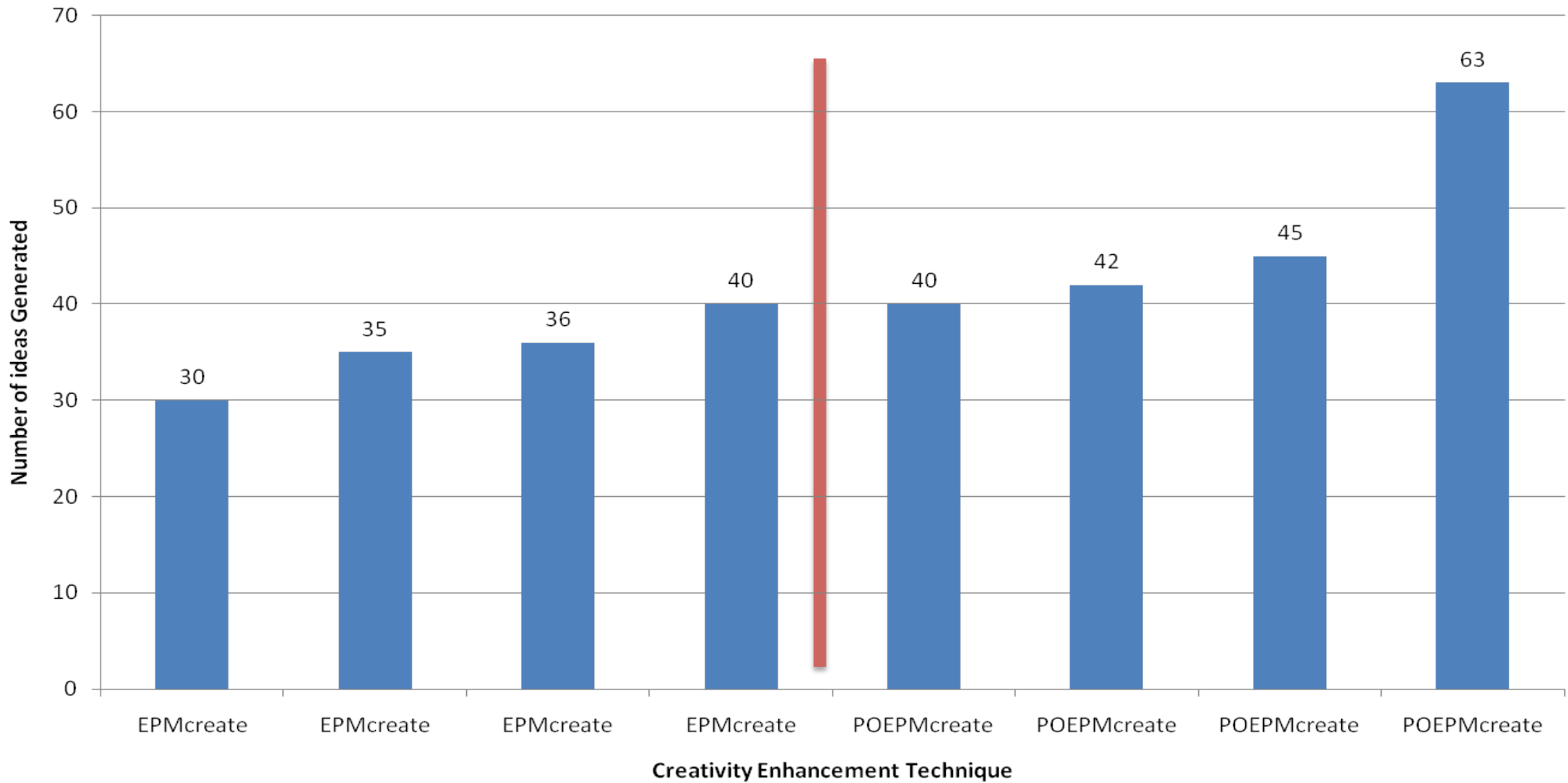
Follow up study (cont.)

A two-sample T-test for unequal variances

POEPMcreate   >   EPMcreate

helps generate

a larger quantity

of requirement ideas than

$\alpha = 0.05$ P $= 0.048$

A two-sample T-test for unequal variances

POEPMcreate   >   EPMcreate

helps generate
more high quality
requirement ideas than

α = 0.05 P = 0.032.

Last SlideS

Q1: Which quality features are addressed by the paper?

A1: Quality of a creativity enhancement technique (CET), POEPMcreate Quality of requirement ideas generated with the help of a CET

Last SlideS (cont.)

Q2: What is the main novelty/contribution of the paper?

A2: An empirical demonstration that as a CET for helping to generate requirement ideas, POEPMcreate is more effective than EPMcreate, which is more effective than brainstorming.

Q3: How will this novelty/contribution improve RE practice or RE research?

A3: It will help make ideas produced in requirements elicitation more creative.

Last SlideS (cont.)

Q4: What are the main problems with the novelty/ contribution and/or with the paper?

A4: The small number of data points in the controlled experiment; replications required

Last SlideS (cont.)

Q5: Can the proposed approach be expected to scale to real-life problems?

A5: Yes! The problem used in the experiment was the enhancement of a real-life Web site and the amount of time used to generate ideas was normal for a real-life requirements elicitation session. The subjects were real people ☺