Survey on Code Smells

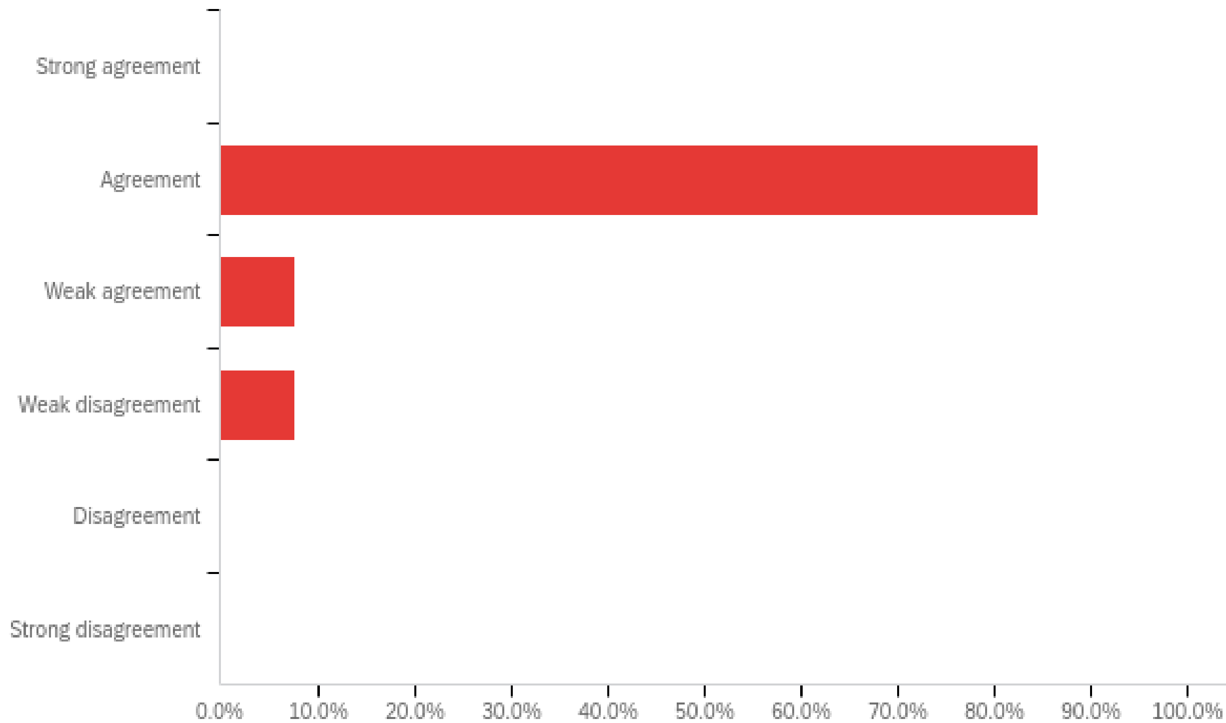
Visualization – Descriptive statistics of responses

Rating scale used

Answer	Rating value
Strong agreement	1
Agreement	2
Weak agreement	3
Weak disagreement	4
Disagreement	5
Strong disagreement	6

Part 2 - Code smells visualization

Q2.2 - SLR FINDING: The vast majority of code smells detection studies do not propose visualization features for their detection.



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	SLR FINDING: The vast majority of code smells detection studies do not propose visualization features for their detection.	2.0	4.0	2.2	0.6	0.3	13

#	SLR FINDING: The vast majority of code smells detection studies do not propose visualization features for their detection.	Percentage
1	Strong agreement	0.0%
2	Agreement	84.6%
3	Weak agreement	7.7%
4	Weak disagreement	7.7%

5	Disagreement	0.0%
6	Strong disagreement	0.0%
	Total	13

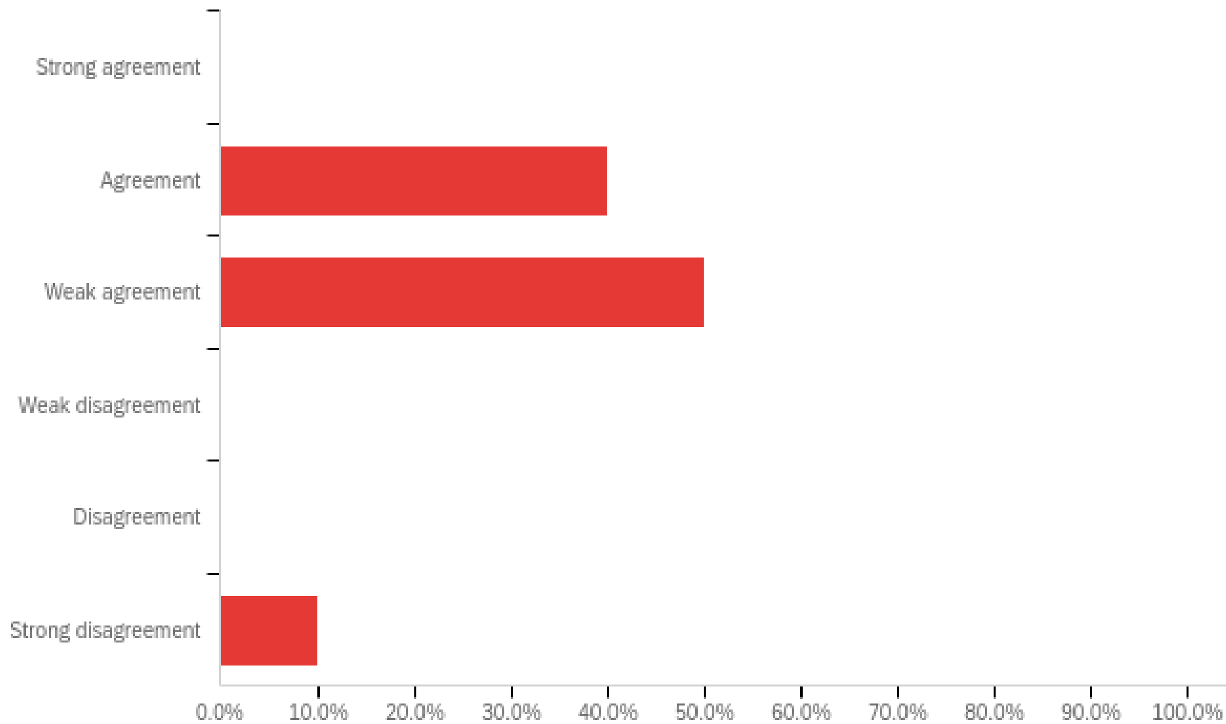
Q2.3 - How do you rate your confidence degree while assessing the previous finding?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How do you rate your confidence degree while assessing the previous finding?	0.0	4.0	2.5	1.2	1.5	12

Q2.4 - Optional justification or comments

Optional justification or comments
Depends what you mean by visualization. Is a simple scatterplot or bar chart also a visualization? If not (i.e. we consider only more 'advanced' visualizations), then I am quite confident about my answer.
I cannot answer the question because I only know few of the respective studies.
I haven't done an SLR.

Q2.5 - SLR FINDING: The vast majority of existing code smells visualization studies did not present evidence of its usage upon large software systems.



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	SLR FINDING: The vast majority of existing code smells visualization studies did not present evidence of its usage upon large software systems.	2.0	6.0	2.9	1.1	1.3	10

#	SLR FINDING: The vast majority of existing code smells visualization studies did not present evidence of its usage upon large software systems.	Percentage
1	Strong agreement	0.0%
2	Agreement	40.0%
3	Weak agreement	50.0%
4	Weak disagreement	0.0%
5	Disagreement	0.0%
6	Strong disagreement	10.0%
	Total	10

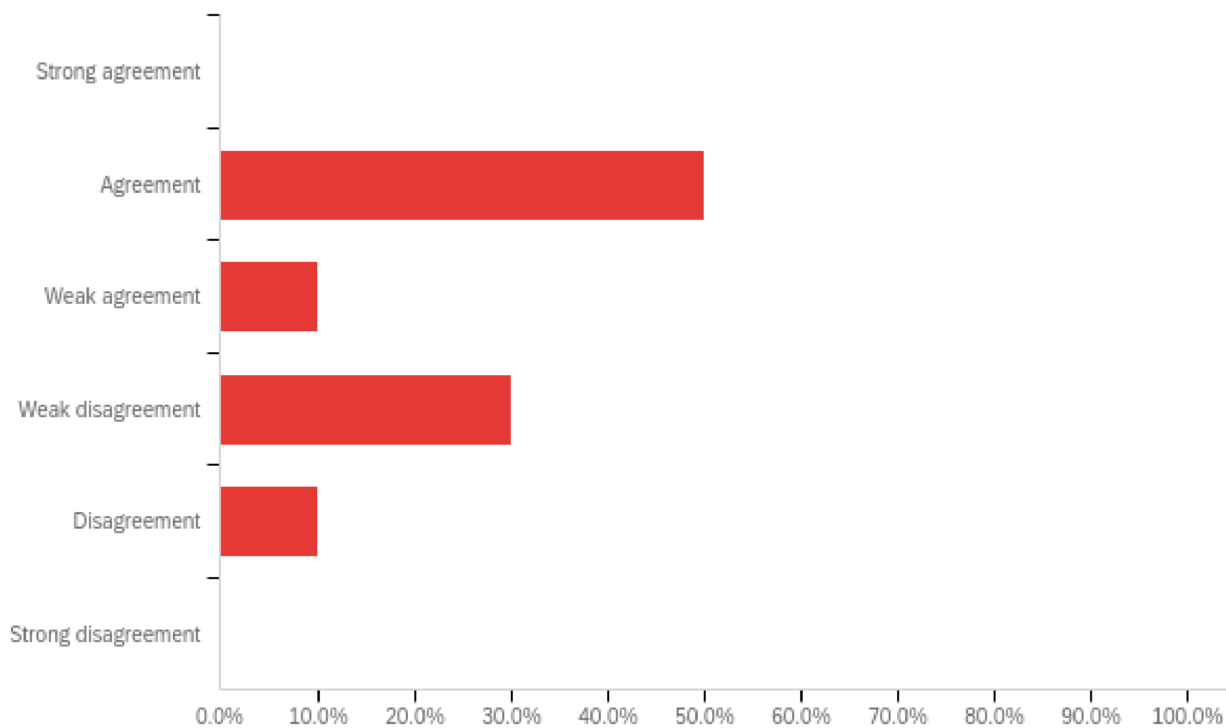
Q2.6 - How do you rate your confidence degree while assessing the previous finding?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How do you rate your confidence degree while assessing the previous finding?	2.0	3.8	2.8	0.5	0.3	8

Q2.7 - Optional justification or comments

Optional justification or comments
Depends what you mean by 'large': 10K lines of code? 100K? More than 1 million? If more than say 200..400K, then I would agree..strongly agree with the question.
I cannot answer the question because I only know few of the respective visualizations..

Q2.8 - SLR FINDING: Software visualization researchers have not adopted specific visualization related taxonomies, such as the ones below, to support the identification of code smells: B. Price, R. Baecker, I. Small, A principled taxonomy of software visualization, Journal of Visual Languages and Computing 4 (3) (1993) 211–266. Roman, G. C., & Cox, K. C. (1993).A taxonomy of program visualization systems. Computer, 26(12), 11-24. Maletic, J. I., Marcus, A., & Collard, M. L. (2002). A task oriented view of software visualization. In Proceedings First International Workshop on Visualizing Software for Understanding and Analysis (pp.32-40). IEEE. Gallagher, K., Hatch, A., & Munro, M. (2008). Software architecture visualization: An evaluation framework and its application. IEEE Transactions on Software Engineering, 34(2), 260-270. Myller, N., Bednarik, R., Sutinen, E., & Ben-Ari, M. (2009). Extending the engagement taxonomy: Software visualization and collaborative learning. ACM Transactions on Computing Education (TOCE), 9(1), 7.



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	SLR FINDING: Software visualization researchers have not adopted specific visualization related taxonomies, such as the ones below, to support the identification of code smells: B. Price, R. Baecker, I. Small, A principled taxonomy of software visualization, Journal of Visual	2.0	5.0	3.0	1.1	1.2	10

	Languages and Computing 4 (3) (1993) 211–266. Roman, G. C., & Cox, K. C. (1993).A taxonomy of program visualization systems. Computer, 26(12), 11-24. Maletic, J. I., Marcus, A., & Collard, M. L. (2002). A task oriented view of software visualization. In Proceedings First International Workshop on Visualizing Software for Understanding and Analysis (pp.32-40). IEEE. Gallagher, K., Hatch, A., & Munro, M. (2008). Software architecture visualization: An evaluation framework and its application. IEEE Transactions on Software Engineering, 34(2), 260-270. Myller, N., Bednarik, R., Sutinen, E., & Ben-Ari, M. (2009). Extending the engagement taxonomy: Software visualization and collaborative learning. ACM Transactions on Computing Education (TOCE), 9(1), 7.						

#	SLR FINDING: Software visualization researchers have not adopted specific visualization related taxonomies, such as the ones below, to support the identification of code smells: B. Price, R. Baecker, I. Small, A principled taxonomy of software visualization, Journal of Visual Languages and Computing 4 (3) (1993) 211–266. Roman, G. C., & Cox, K. C. (1993).A taxonomy of program visualization systems. Computer, 26(12), 11-24. Maletic, J. I., Marcus, A., & Collard, M. L. (2002). A task oriented view of software visualization. In Proceedings First International Workshop on Visualizing Software for Understanding and Analysis (pp.32-40). IEEE. Gallagher, K., Hatch, A., & Munro, M. (2008). Software architecture visualization: An evaluation framework and its application. IEEE Transactions on Software Engineering, 34(2), 260-270. Myller, N., Bednarik, R., Sutinen, E., & Ben-Ari, M. (2009). Extending the engagement taxonomy: Software visualization and collaborative learning. ACM Transactions on Computing Education (TOCE), 9(1), 7.	Percentage
1	Strong agreement	0.0%
2	Agreement	50.0%
3	Weak agreement	10.0%
4	Weak disagreement	30.0%
5	Disagreement	10.0%
6	Strong disagreement	0.0%
	Total	10

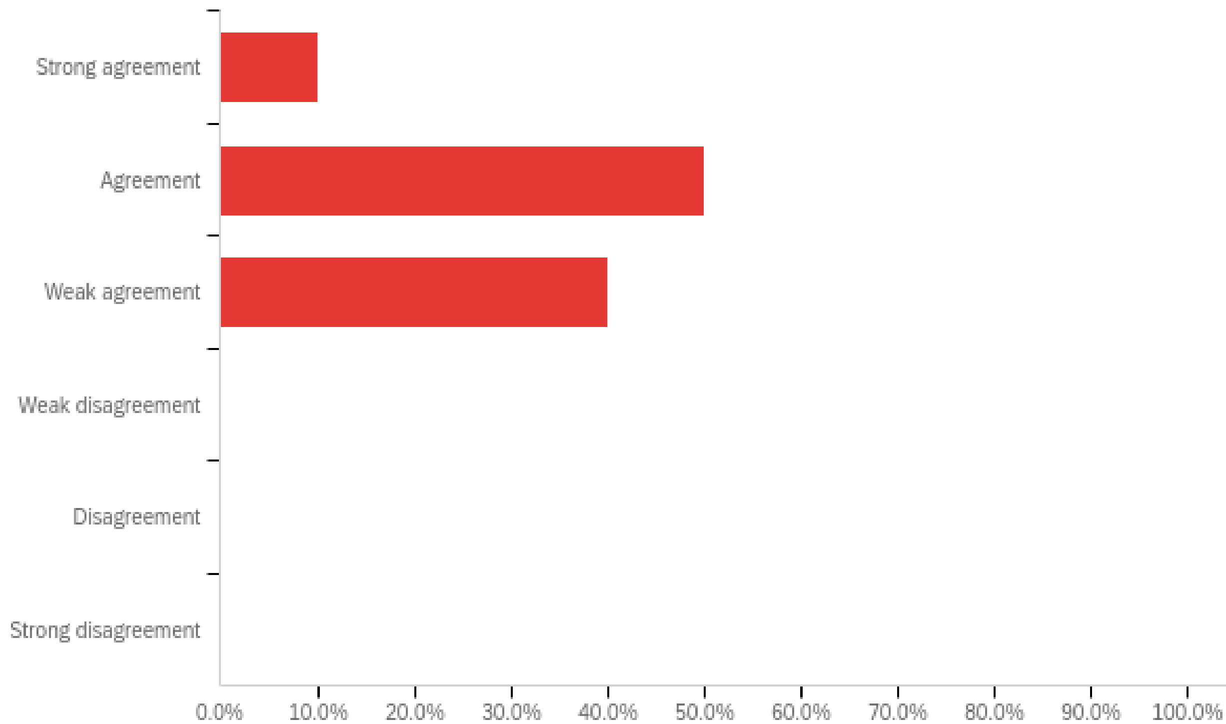
Q2.9 - How do you rate your confidence degree while assessing the previous finding?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How do you rate your confidence degree while assessing the previous finding?	0.0	3.7	2.2	1.2	1.4	7

Q2.10 - Optional justification or comments

Optional justification or comments
I know of quite a large number of software vis papers (and researchers) who have adopted at least one of the above mentioned taxonomies.
Not convinced that the referenced papers are (still) relevant. I know the one of Maletic et al. - it's okay, but does not go into details. I don't use it for my own papers.

Q2.11 - OPINION: If visualization related taxonomies were used in the implementation of code smells detection tools, that could enhance their effectiveness.



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	OPINION: If visualization related taxonomies were used in the implementation of code smells detection tools, that could enhance their effectiveness.	1.0	3.0	2.3	0.6	0.4	10

#	OPINION: If visualization related taxonomies were used in the implementation of code smells detection tools, that could enhance their effectiveness.	Percentage
1	Strong agreement	10.0%
2	Agreement	50.0%
3	Weak agreement	40.0%
4	Weak disagreement	0.0%
5	Disagreement	0.0%
6	Strong disagreement	0.0%
	Total	10

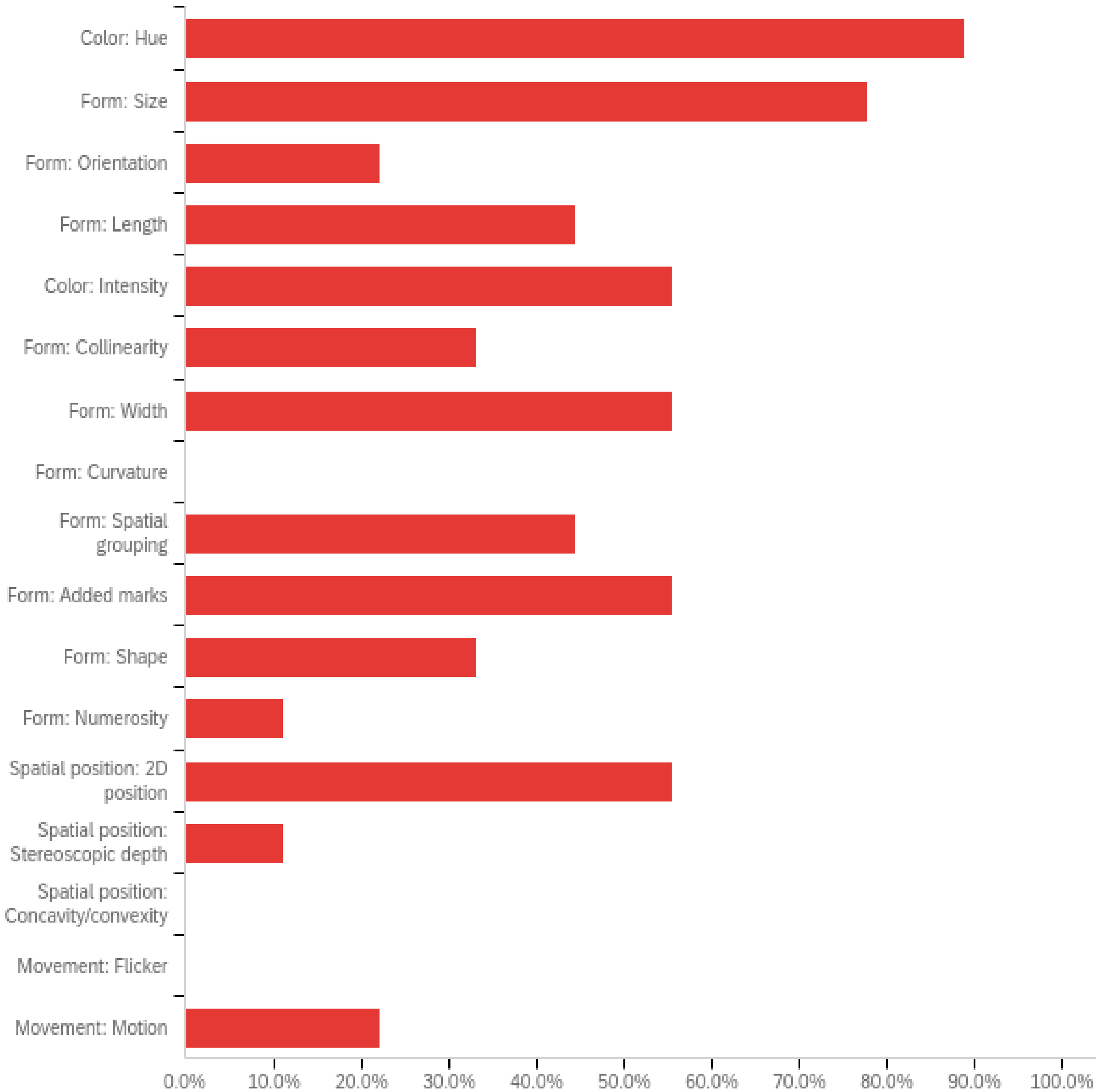
Q2.12 - How do you rate your confidence degree while expressing the previous opinion?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How do you rate your confidence degree while expressing the previous opinion?	1.2	4.0	3.0	0.9	0.8	6

Q2.13 - Optional justification or comments

Optional justification or comments
I'm not sure that a taxonomy is directly contributing to an implementation as such. A taxonomy helps to organize related work, ideas, etc, and present/frame one's own work, but doesn't help directly when implementing a new method.
Taxonomies, in general, can help to compare approaches, not sure they are that helpful for developing tools/approaches.

Q2.14 - OPINION: Which of the following visual attributes have you implemented in tools targeting the support of code smells identification? Consider bellow resources discussed in the literature [Mazza, R. (2009). Introduction to information visualization. Springer Science & Business Media.]:



#	OPINION: Which of the following visual attributes have you implemented in tools targeting the support of code smells identification? Consider bellow resources discussed in the literature [Mazza, R. (2009). Introduction to information visualization. Springer Science & Business Media.]:	Percentage
1	Color: Hue	14.5%
8	Form: Size	12.7%

3	Form: Orientation	3.6%
4	Form: Length	7.3%
22	Color: Intensity	9.1%
2	Form: Collinearity	5.5%
9	Form: Width	9.1%
23	Form: Curvature	0.0%
24	Form: Spatial grouping	7.3%
25	Form: Added marks	9.1%
26	Form: Shape	5.5%
27	Form: Numerosity	1.8%
28	Spatial position: 2D position	9.1%
29	Spatial position: Stereoscopic depth	1.8%
30	Spatial position: Concavity/convexity	0.0%
31	Movement: Flicker	0.0%
32	Movement: Motion	3.6%
	Total	55

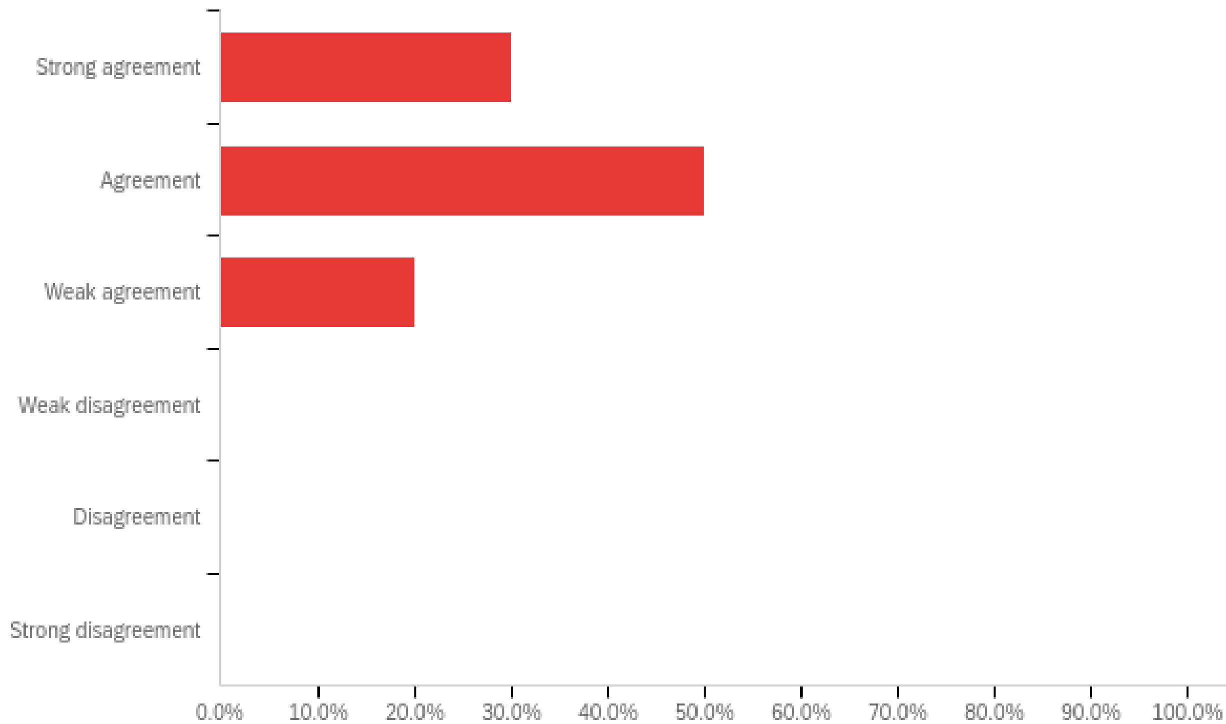
Q2.15 - How do you rate your confidence degree while expressing the previous opinion?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How do you rate your confidence degree while expressing the previous opinion?	3.0	4.0	3.6	0.4	0.1	8

Q2.16 - Optional justification or comments

Optional justification or comments
I have co-authored one publication in this area using Parallel coordinates and scatterplots. Of course, these visualization relate to the above encodings, but it's not clear in all cases how. The above encodings seem a bit too low-level to make an appropriate classification here.
how on earth is this clearly factual question an opinion about which I can have a confidence

Q2.17 - OPINION: The combined use of collaboration (among software developers) and visual resources may increase the effectiveness of code smells detection.



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	OPINION: The combined use of collaboration (among software developers) and visual resources may increase the effectiveness of code smells detection.	1.0	3.0	1.9	0.7	0.5	10

#	OPINION: The combined use of collaboration (among software developers) and visual resources may increase the effectiveness of code smells detection.	Percentage
1	Strong agreement	30.0%
2	Agreement	50.0%
3	Weak agreement	20.0%
4	Weak disagreement	0.0%
5	Disagreement	0.0%
6	Strong disagreement	0.0%
	Total	10

Q2.18 - How do you rate your confidence degree while expressing the previous opinion?

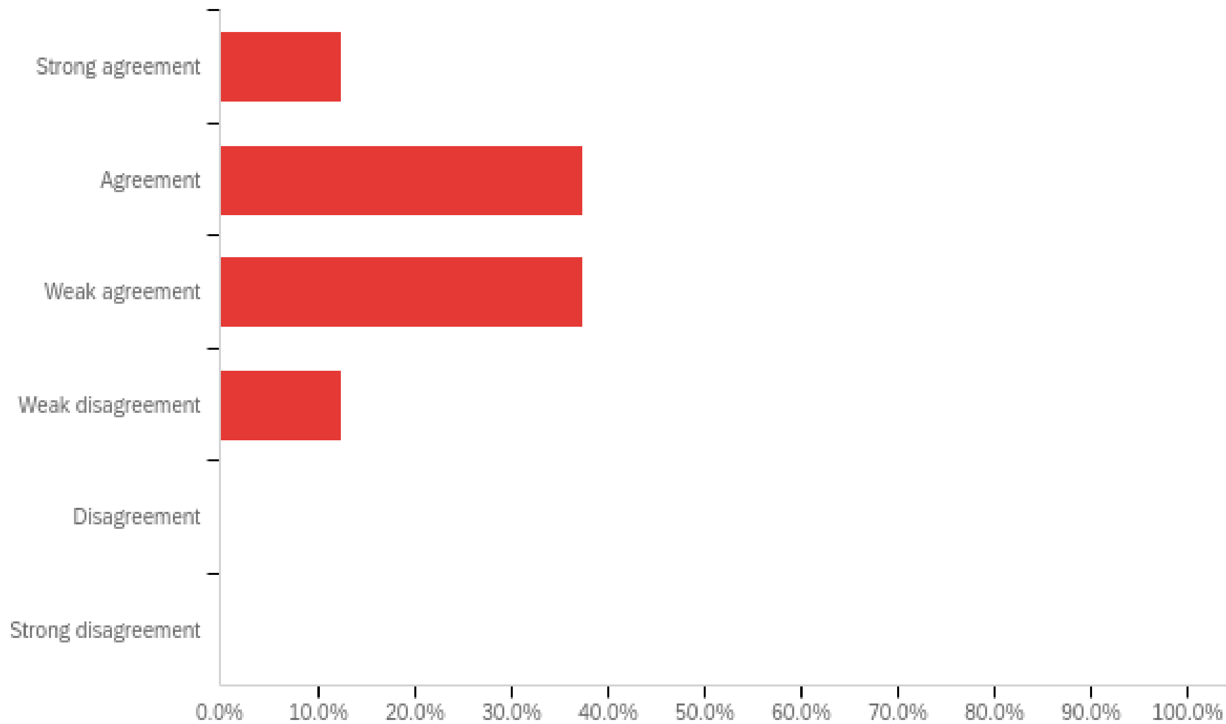
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How do you rate your confidence degree while expressing the previous opinion?	3.0	4.0	3.8	0.3	0.1	7

Q2.19 - Optional justification or comments

Optional justification or comments
Not sure I understand the question.

Part 3 - Code smells detection techniques

Q3.2 - SLR FINDING: The most frequently used code smells detection techniques are based on rule-based approaches.



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	SLR FINDING: The most frequently used code smells detection techniques are based on rule-based approaches.	1.0	4.0	2.5	0.9	0.8	8

#	SLR FINDING: The most frequently used code smells detection techniques are based on rule-based approaches.	Percentage
1	Strong agreement	12.5%
2	Agreement	37.5%
3	Weak agreement	37.5%
4	Weak disagreement	12.5%
5	Disagreement	0.0%

6	Strong disagreement	0.0%
	Total	8

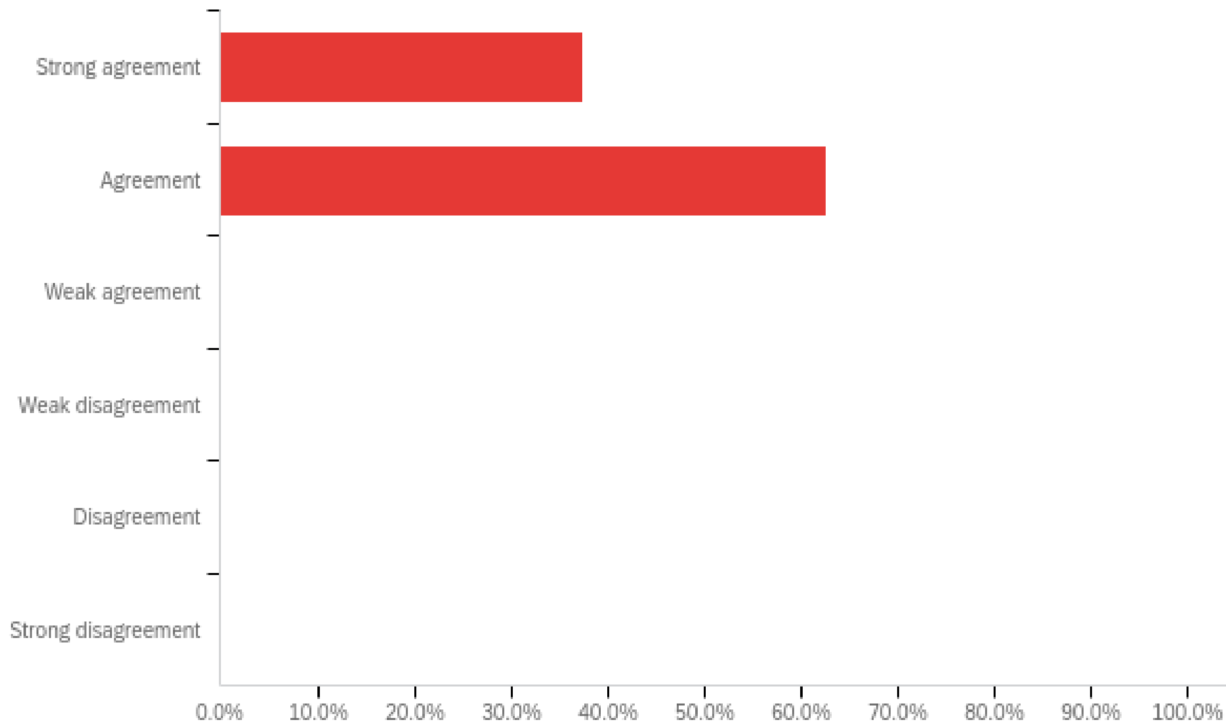
Q3.3 - How do you rate your confidence degree while assessing the previous finding?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How do you rate your confidence degree while assessing the previous finding?	0.8	3.6	2.3	0.9	0.9	8

Q3.4 - Optional justification or comments

Optional justification or comments
Sounds plausible, but I don't have a good overview on "frequently used code smell detection techniques"
parsing and analysis (using algorithms) of metamodels are important

Q3.5 - SLR FINDING: Very few code smells detection studies provide their oracles (a tagged dataset for training detection algorithms).



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	SLR FINDING: Very few code smells detection studies provide their oracles (a tagged dataset for training detection algorithms).	1.0	2.0	1.6	0.5	0.2	8

#	SLR FINDING: Very few code smells detection studies provide their oracles (a tagged dataset for training detection algorithms).	Percentage
1	Strong agreement	37.5%
2	Agreement	62.5%
3	Weak agreement	0.0%
4	Weak disagreement	0.0%
5	Disagreement	0.0%
6	Strong disagreement	0.0%
	Total	8

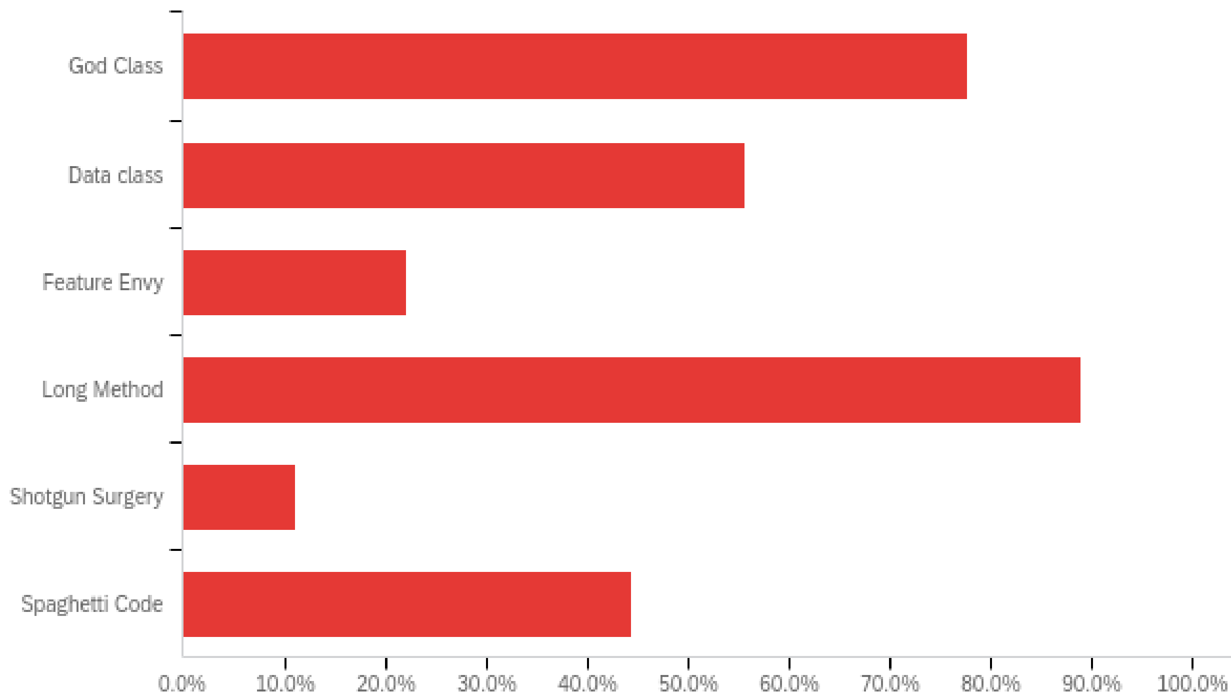
Q3.6 - How do you rate your confidence degree while assessing the previous finding?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How do you rate your confidence degree while assessing the previous finding?	1.3	4.0	3.2	0.8	0.7	7

Q3.7 - Optional justification or comments

Optional justification or comments
This is partly inherent, since not all code smell detection algorithms work with a 'training' phase as such (like in supervised machine learning).
Don't know

Q3.8 - OPINION: Please select the 3 most often detected code smells.



Data source misconfigured for this visualization

#	OPINION: Please select the 3 most often detected code smells.	Percentage
1	God Class	25.9%
2	Data class	18.5%
3	Feature Envy	7.4%
4	Long Method	29.6%
5	Shotgun Surgery	3.7%
6	Spaghetti Code	14.8%
	Total	27

Q3.9 - How do you rate your confidence degree while expressing the previous opinion?

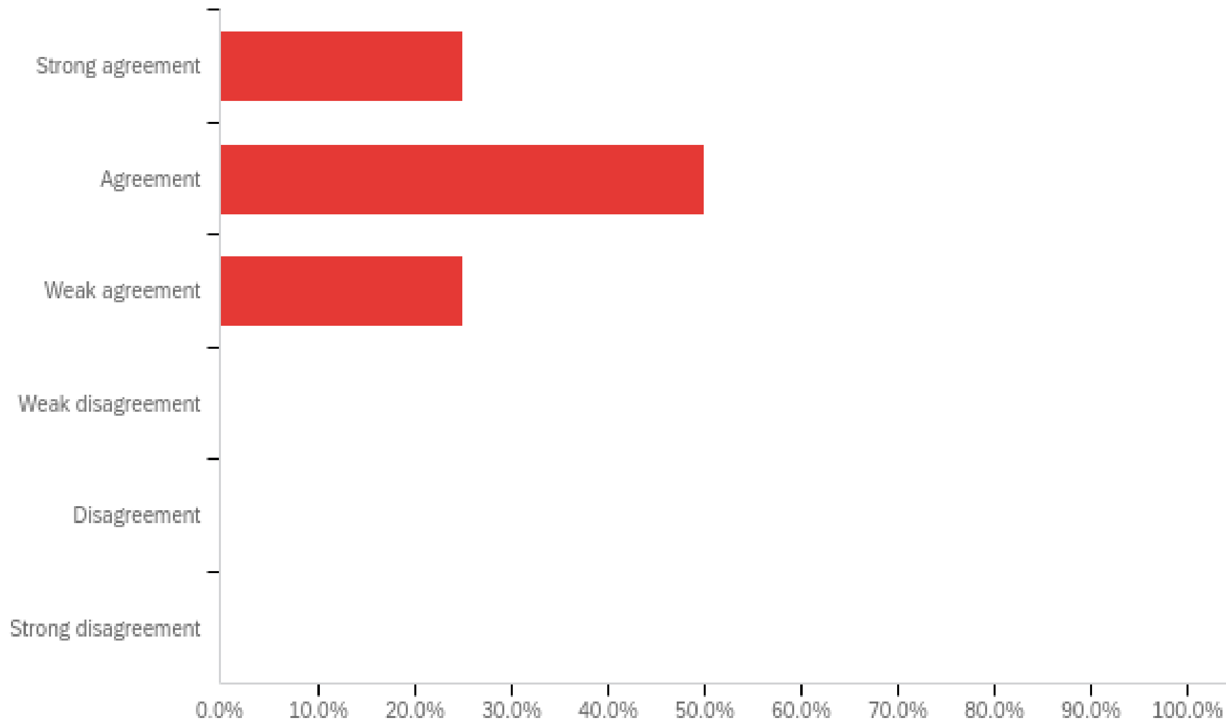
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How do you rate your confidence degree while expressing the previous opinion?	1.1	4.0	3.0	1.1	1.2	6

Q3.10 - Optional justification or comments

Optional justification or comments
Don't know

Part 4 - Code smells detection effectiveness

Q4.1 - SLR FINDING: In the detection of simpler code smells (e.g. Long Method or God Class), the achieved precision and recall of detection techniques can be very high (up to 100%).



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	SLR FINDING: In the detection of simpler code smells (e.g. Long Method or God Class), the achieved precision and recall of detection techniques can be very high (up to 100%).	1.0	3.0	2.0	0.7	0.5	8

#	SLR FINDING: In the detection of simpler code smells (e.g. Long Method or God Class), the achieved precision and recall of detection techniques can be very high (up to 100%).	Percentage
1	Strong agreement	25.0%
2	Agreement	50.0%
3	Weak agreement	25.0%

4	Weak disagreement	0.0%
5	Disagreement	0.0%
6	Strong disagreement	0.0%
	Total	8

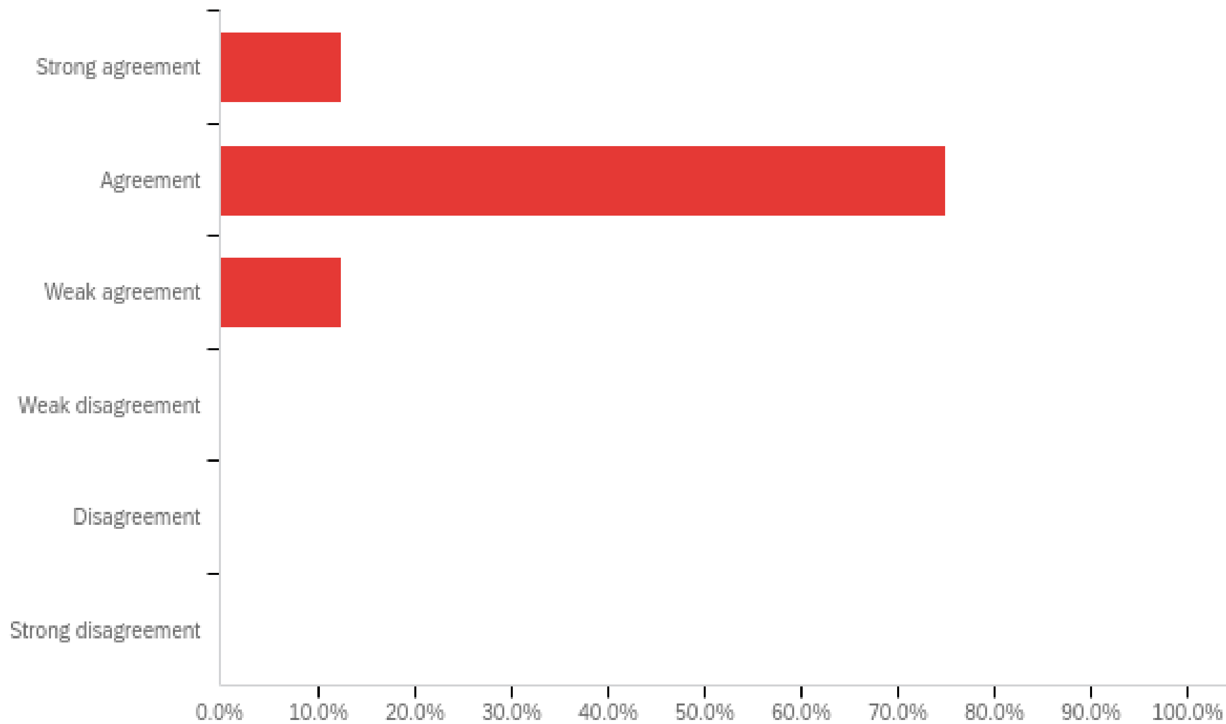
Q4.2 - How do you rate your confidence degree while assessing the previous finding?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How do you rate your confidence degree while assessing the previous finding?	2.3	4.0	3.3	0.6	0.3	7

Q4.3 - Optional justification or comments

Optional justification or comments
Depends a bit on details e.g. the programming language we consider; for e.g. Java or C#, such things work far better/more precisely than e.g. when one aims to support the _entire_ C++ standard (which is huge).
Don't know

Q4.4 - SLR FINDING: When the complexity of code smells is greater (e.g. Divergent Change or Shotgun Surgery), the precision and recall in detection are much lower than in simpler code smells.



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	SLR FINDING: When the complexity of code smells is greater (e.g. Divergent Change or Shotgun Surgery), the precision and recall in detection are much lower than in simpler code smells.	1.0	3.0	2.0	0.5	0.3	8

#	SLR FINDING: When the complexity of code smells is greater (e.g. Divergent Change or Shotgun Surgery), the precision and recall in detection are much lower than in simpler code smells.	Percentage
1	Strong agreement	12.5%
2	Agreement	75.0%
3	Weak agreement	12.5%
4	Weak disagreement	0.0%
5	Disagreement	0.0%

6	Strong disagreement	0.0%
	Total	8

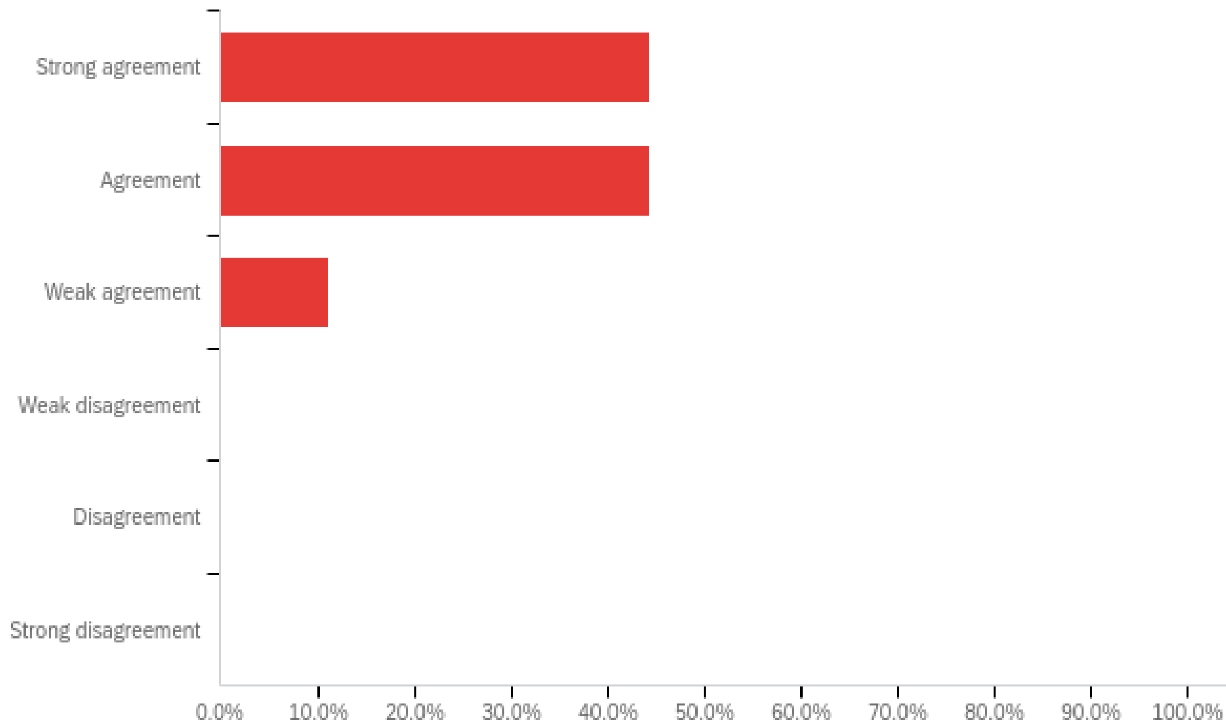
Q4.5 - How do you rate your confidence degree while assessing the previous finding?

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How do you rate your confidence degree while assessing the previous finding?	1.1	4.0	2.8	1.0	1.0	7

Q4.6 - Optional justification or comments

Optional justification or comments
Sure; but of course it depends what you mean by 'much smaller'.
Don't know

Q4.7 - SLR FINDING: There are few oracles (a tagged dataset for training detection algorithms) shared and publicly available. The existence of shared and collaborative oracles could improve the state of the art in code smells detection research.



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	SLR FINDING: There are few oracles (a tagged dataset for training detection algorithms) shared and publicly available. The existence of shared and collaborative oracles could improve the state of the art in code smells detection research.	1.0	3.0	1.7	0.7	0.4	9

#	SLR FINDING: There are few oracles (a tagged dataset for training detection algorithms) shared and publicly available. The existence of shared and collaborative oracles could improve the state of the art in code smells detection research.	Percentage
1	Strong agreement	44.4%
2	Agreement	44.4%
3	Weak agreement	11.1%
4	Weak disagreement	0.0%

5	Disagreement	0.0%
6	Strong disagreement	0.0%
	Total	9

Q4.8 - How do you rate your confidence degree while assessing the previous finding?

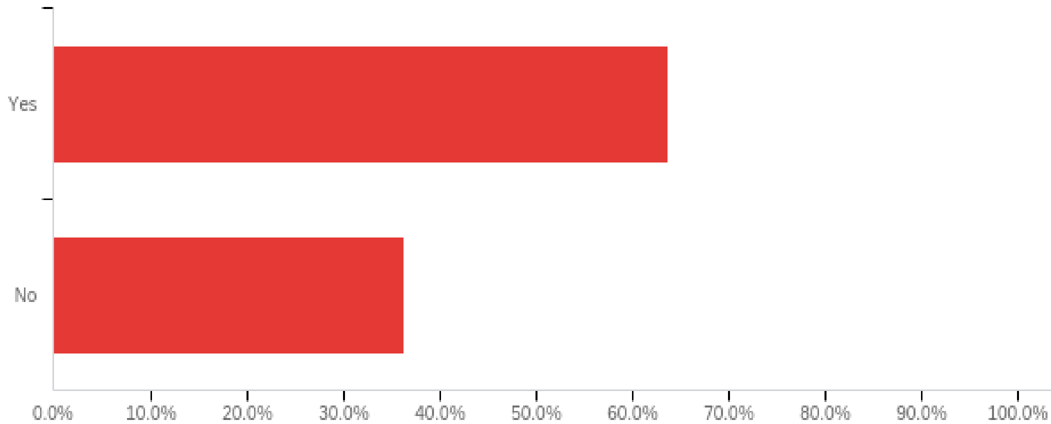
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How do you rate your confidence degree while assessing the previous finding?	3.0	4.0	3.6	0.4	0.2	6

Q4.9 - Optional justification or comments

Optional justification or comments
Don't know

Part 5 - Respondents' info

Q5.1 - OPTIONAL   Do you want to have first-hand access to the Systematic Literature Review?



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	OPTIONAL Do you want to have first-hand access to the Systematic Literature Review?	1.0	2.0	1.4	0.5	0.2	11

#	OPTIONAL Do you want to have first-hand access to the Systematic Literature Review?	Percentage
1	Yes	63.6%
2	No	36.4%
	Total	11

Q5.2 - Please provide your Email address below.   NOTE: You will not be identified in any report that is produced using the information you have provided in this questionnaire and your email will not be used for any other purpose, except for sending you the SLR.

Email address
****