

Article

An example of a science communication evaluation study: Discovery07, a Dutch science party

Cees M. Koolstra

This article presents an example of how a public science party was evaluated. The main goals of the science party, to increase the positive image of science and present an attractive science event, were evaluated in two ways. First, web surveys were used to determine the image of science before and after the event among paying visitors, invited guests, and a control group (N = 149). Second, during the event, visitors were interviewed about their experiences at the event (N = 124). The survey study showed that the image of science was very positive among all three groups of respondents. As no differences were found between pre- and post-tests, participation in the event did not lead to a more positive image of science. The results of the interviews suggested that visitors highly appreciated the event. In the Discussion, the evaluation study is analyzed and possibilities/limitations for future general use are discussed.

Introduction

Discovery07 was organized on September 27 of 2007 in Amsterdam, the Netherlands. It was a public event that combined science with art activities in the form of a party. The party was held at the terrain of the Westergasfabriek of which three locations (or halls) were used.

The research reported here pertains to an evaluation study of Discovery07. The reason for the publication of this report is that it may be assumed that there is an increasing interest among SC (science communication) professionals to gain insight in how public SC events can be evaluated. In many cases, for example when the European Union subsidizes SC events, evaluation of the effects of the event is obligatory. In addition, evaluation studies of SC events may help organizers to get new ideas of how to adjust specific elements or activities of the event to make them more effective. It may be noted that the evaluation of effects can pertain to very different goals. In some cases a goal might be to increase the public knowledge of science, whereas in other cases a goal might be to give visitors a pleasant time. The present evaluation study concerned the possible effects of increasing the positive image of science and scientists and providing the public with an attractive science event through which they would become interested in the combination of science with art.

The idea of reporting the present study here is certainly *not* to provide readers with a (nearly) perfect example of how an evaluation study should be conducted or how the visit of this specific SC event led to large effects. The main idea of the report is to present an actual example of a SC evaluation study in which some possibilities as well as some limitations of evaluations are shown. On the one hand, the report demonstrates some “good” principles of using a pre- and post-test, control group, and well-developed measuring instruments, while on the other hand, not unknown to most research in the social sciences, the report illustrates limitations pertaining to measurement, analyses, and conclusions. Discussion of the possibilities and limitations of this exemplar evaluation study hopefully provides readers with the insight of how in the future valid evaluations of SC activities can be developed.

Discovery07

The event was announced as “a party with a meaning” combining live music, live activities of scientific research, DJ's, VJ's, cocktails, movies, performances, experiments, visuals, games, and presentations. The activities pertaining to science were all performed by young scientists. The event was organized by the *Stichting Very Disco*, a network of young scientists who are “as wildly enthusiastic about new science as about music, art and new ideas” (see <http://www.discovery07.nl>).

Goals of Discovery07

Discovery07 had three informal goals. First, the event was aimed at showing to a young, general public that science can be a fascinating endeavor. The second goal was to show that there are many young and creative scientists. The third goal was to inspire young scientists and science professionals to give a nontraditional interpretation of making science popular among the young public. The two formal goals were to increase the public's positive image of science and scientists and to offer the young public an attractive event in which science and art was combined.

Design of the evaluation study

The evaluation study was aimed at determining whether the two formal goals would be achieved: Would visiting the event lead to a more positive image of science and scientists and how would the visitors experience and judge the event? The first goal was investigated by using a pre-test and post-test that measured the image of science and scientists among visitors. Because the same test was also conducted among a small group of people who did not visit the event, there was the possibility of comparing the visitors' data with a control group. The pre-test was conducted among future visitors who had ordered tickets for the event by sending them a request to participate in the study through a "mass e-mail" invitation. Pre- and post-test were identical with the exception that the post-test also included additional questions about whether respondents had visited Discovery07 and about whether they were exposed to other SC experiences in the time period between pre- and post-test such as watching science television programs, reading books or magazines about science and visiting other science events (such as science festivals or lectures presented by scientists). These additional questions provided the opportunity to control for the possible confounding influence of these experiences in the analyses.

Asking about background characteristics, such as age, gender, educational level of respondents in pre- and post-tests provided the possibility to determine whether visitors of the event could be compared with their peers in general and whether possible effects would differ between subgroups divided on the basis of these background characteristics. It would be possible, for example, to analyze whether visiting the event would be more (or less) effective for female or male visitors or for younger or older visitors. The post-test was conducted about one week after the occurrence of the event. The same measurement was also employed to a control group of persons who did not visit the event.

Measurement instruments

The image of science and scientists was measured on the basis of two existing instruments. First, the ISSS (*Image of Science and Scientists Scale*), originally developed by Krajcovich and Smith [4] was used. A recent study on the reliability and validity of this instrument conducted by Marshall et al. [5] showed that the short version of the instrument with 17 items was amply adequate. The instrument included statements such as "Science is responsible for progress", "A scientist's work is dull" and "I would like to be a scientist". For each statement respondents could indicate their level of agreement on a 5-point Likert response scale varying from "Strongly agree" to "Strongly disagree". The second measurement was a set of 5 statements widely used to measure the image of science and scientists (sometimes with slight variations in the wording of the statements) in studies such as the *Eurobarometer* research in Europe and *Science and Engineering Indicators* conducted by the *National Science Foundation* in the U.S. The same instrument was also used in an earlier Dutch study on the image of science conducted by Becker and Van Rooijen [1] among the public, science journalists, and members of parliament. Examples of statements are "The benefits of science are greater than the harmful effects it may have" and "We depend too much on science and not enough on faith (and our emotions)". Respondents could indicate their level of agreement with the statements on a 5-point Likert scale varying from "Strongly agree" to "Strongly disagree". The advantage of using existing measurement instruments was that the reliability of the instruments was already established (and adequate) and that the average scores that would be found in the present study could be compared with those found in earlier studies.

An additional component of the data collection was a measurement among a control group of people who were planning to visit Discovery07 but actually did not come to the event and people who in the past had showed interest in other SC activities presented by the organizers of Discovery07. All of these respondents were selected from a large database build by the Discovery07 organization.

	Pre-test (before the visit of Discovery07)	Measurement during Discovery07	Post-test (a week after the visit of Discovery07)
Research sample	People who had reserved tickets	Visitors	Visitors and control group
Measured variables	Image of science and scientists (and background characteristics)	Appreciation of the event including its different activities (and background characteristics)	Image of science and scientists (and background characteristics and recent SC experiences)
Method	Web survey	Oral interviews	Web survey

Table 1. Design of the study with three types of measurements.

In addition to the pre- and post-tests, short face-to-face interviews were conducted during the event among a sample of visitors. As about 1000 to 1500 were expected to participate in the event, a sample of about 10% of the visitors was interviewed. These interviews were aimed at determining whether visitors liked the event, which activities in the event were appreciated most (and least), and what people thought about the combination of science and art. Near the end of the interview, people were also asked to assign a (school-type) grade between 1 and 10 to the event, with 1 as the lowest and 10 as the highest possible judgment. The interviews were conducted by 3 trained senior students who registered the respondents' answers by writing them down on paper.

Table 1 presents an overview of the design of the evaluation study with the three types of measurements (before, during, and after the event), the different research samples, and the measured variables included in the study.

In total, 149 respondents participated in the pre- and post-test administered by means of the web surveys. The reliability of the 17-item image of science scale developed by Krajcovich and Smith [4] was sufficient: Cronbach's alpha was .78. The reliability of the 5-item Eurobarometer image of science scale was somewhat lower but still reasonable: Cronbach's alpha was .69. For the measurement during Discovery07, interviews were held with 124 visitors during the event.

Procedures

The pre-test was employed among persons who had ordered tickets for the event through the internet by sending them an e-mail invitation to participate in a web survey one day before the event took place. The invitation was also sent to people who were on a guest list. The website through which respondents could participate in the pre-test web survey was closed a few hours before the event took place. Ultimately, on the day that the website was available, 48 persons responded to the invitation by filling in the web survey. One week after the event was held, another mass e-mail invitation was sent to people in the database of the Discovery07 organization. This database consisted of e-mail addresses of: persons who had bought entrance tickets for Discovery07, persons who were on the guest list (e.g., scientists and artists with some connection to the organization), and a group of people who had visited other SC events presented by the Discovery07 organization. 101 respondents participated in the post-test survey, of which 58 persons had actually visited Discovery07 and 43 persons who did not visit the event. This last group was considered as the control group of persons who were highly comparable to the group of visitors. A check to compare respondents' identities in the pre- and post-test showed that there were no people who participated in pre- as well as post-test. Therefore, formally the design of this part of the study was quasi-experimental with separate pre-post samples.

The interviews among visitors during the event were conducted by 3 trained senior students. Visitors were randomly selected and asked to participate in the interview, and each interview included a short set of open questions asked in each interview. Then, the respondent was asked to assign a grade to the event. Additional questions asked the respondent's age and highest educational level. Answers were written down (in short notation).

	Pre-test	Post-test	Post-test among respondents who were not present
Visitors	68.00 ($N = 21$)	69.36 ($N = 11$)	68.70 ($N = 10$)
Guests	68.42 ($N = 26$)	68.66 ($N = 47$)	67.64 ($N = 33$)
<i>M</i> (mean)	68.23 ($N = 47$)	68.79 ($N = 58$)	67.88 ($N = 43$)

Table 2. Comparisons of average scores on the 17-item image of science scale for visitors, guests (and non-visitors) in the pre- and post-test.

	Pre-test	Post-test	Post-test among respondents who were not present
Visitors	20.19 ($N = 21$)	20.45 ($N = 11$)	20.90 ($N = 10$)
Guests	20.12 ($N = 25$)	20.96 ($N = 47$)	20.24 ($N = 33$)
<i>M</i> (mean)	20.15 ($N = 46$)	20.86 ($N = 58$)	20.40 ($N = 43$)

Table 3. Comparisons of average scores on the 5-item image of science scale for visitors, guests (and non-visitors) in the pre- and post-test.

Results

Pre- and post-tests using the web surveys

Of the 149 respondents who participated in the pre- and post-tests, 45% was female and 55% male. 30% of the respondents was between the age of 26 and 30, 24% was between 20 and 25 years old, 23% older than 36, and 16% was between 31 and 35 years old. The highest educational level was university for 79% of the sample; 14% was at the high professional level and 7% at the high general level.

At the pre-test, the average total score on the 17-item image of science scale (ISSS) was 68.23 ($N = 47$). This means that the average item score was 4.01 (minimum was 1 and maximum 5) indicating that the respondents had a very positive image of science and scientists. The highest item score was found for the statement "Science is responsible for progress" ($M = 4.34$) with which respondents agreed most as compared with other statements. The statement "A scientist has little social life" ($M = 3.50$) received the (relatively) lowest score which indicated that a relatively small number of respondents did agree with the statement.

The average total score on the pre-test of the 5-item image of science scale (Eurobarometer) was 20.15 ($N = 46$) which means that the average item score for this scale was also 4.00 (minimum 1 and maximum 5). The highest average item score was found for the statement "The benefits of science are greater than the harmful effects it may have" ($M = 4.22$) which indicated that respondents mostly agreed with this statement.

At the post-test the average total score on the 17-item image scale (ISSS) was 68.79 ($N = 58$) among respondents who were present at Discovery07. Visitors scored somewhat higher ($M = 69.36$) than guests ($M = 68.66$), but this difference was not statistically significant ($p < .05$). The average score at the post-test of the control group of respondents who did not visit the event was $M = 67.88$ ($N = 43$).

The post-test average total score of the 5-item image scale was $M = 20.86$ for respondents who visited Discovery07 and $M = 20.40$ for those who were not present at the event.

Tables 2 and 3 present all the average scores for both the 17-item and 5-item image scales. Statistical tests for establishing differences between pre- and post-test, visitors and guests, and respondents who were present or not at the event were conducted, but they are not fully reported here, because none of the comparisons showed significant differences (at $p < .05$). Therefore, the conclusion was that for both image scales there were no statistically significant differences between the various subgroups who participated in the web surveys.

Extra analyses were conducted to investigate whether there were differences in image scores (on both scales) between subgroups divided on the basis of gender, age, and educational level. There were no differences for gender and age, but there were differences for educational level. Respondents with the highest educational level (university) had significantly higher scores on both image scales than respondents with (somewhat) lower educational levels (medium and high professional education and high general education): respondents with an academic educational level scored respectively $M = 69.16$ and

	Respondents present study	Public	Science journalists	Members of parliament
In my daily life, it is not important to know about science (disagree).	92%	50%	--*	--*
We depend too much on science and not enough on faith (and our emotions) (disagree).	18%	35%	16%	39%
Because of their knowledge, scientists have a power that makes them dangerous (agree).	13%	56%	31%	36%
The harmful effects of science are greater than the benefits it may have (disagree).	88%	63%	88%	77%

Table 4. Comparisons of answers to 4 image of science questions asked in the present study with those of the Dutch public, science journalists, and members of parliament in the study conducted by Becker and Van Rooijen [1].

*Note. In Becker and Van Rooijen [1] the first question was not asked to science journalists and members of parliament.

$M = 20.92$ on the 17- and 5-item scales, whereas those with lower educational level scored $M = 65.72$ and $M = 19.10$ (p 's < .05). The analyses to investigate whether participation in other SC activities (watching television programs, reading books/magazines, and visiting events pertaining to science) affected the scores showed that these experiences did not influence the scores. Therefore, correcting and/or controlling the scores for these activities was not necessary.

The average scores on the 17-item image of science scale found in the present study were clearly higher than those found the American study conducted by Krajcovich and Smith [4], although the comparison is somewhat problematic as the American study used the older version of the instrument with more items. Unfortunately, the recent American study conducted by Marshall et al. [5] did not report mean scores so that a comparison with that study is impossible.

It must be noted that there were some respondents who gave negative comments on the 17-item image of science questionnaire (at the end of the survey where people were provided the possibility to give a reaction on the survey): 4 respondents thought that the statements were too "stereotypical"; 3 other respondents found the statements too "one-sided" and/or "negative"; 2 people complained that some of the statements were "old-fashioned"; and 1 respondent found the statements too "suggestive". 2 of these respondents mentioned specific items as negative examples: "A scientist neglects his/her family" (stereotypical) and "A scientist works in a dreary laboratory" (old-fashioned).

With regard to the 5-item scale a comparison with the scores found in the Dutch study conducted by Becker and Van Rooijen [1] it must be noted that their study condensed the 5 possible answering categories into 3 categories: (fully) agree, neither agree nor disagree, and (fully) disagree. On the level of individual items, for 4 out of 5 items table 4 shows to what degree the answers of respondents in the present study compare to those in the study conducted by Becker and Van Rooijen [1]. Interestingly, Becker and Van Rooijen reported not only scores for the public (for 3 out of 4 items), but also for science journalists and members of parliament.

The comparisons show that respondents in the present study have a much more positive image of science and scientists than the general Dutch public. An example is the statement "In my daily life, it is not important to know about science" with which 92% of the respondents of the present study and only 50% of the Dutch public disagreed. In the present study only 13% of the respondents agreed with the statement "Because of their knowledge, scientists have a power that makes them dangerous", whereas in the earlier study 56% of the public, 31% of the science journalists, and 36% of the parliament members agreed with this (negative) statement. In this example the respondents in the present study were not only much less negative about the "danger" of scientists than the public, but even less negative than science journalists.

Duration of presence	Grade	Number of respondents
0-1 hour	7.15	13
1-2 hours	7.62	42
2-3 hours	7.62	37
3-4 hours	7.73	22
4-5 hours	8.33	6
<i>M</i>	7.63	120

Table 5. Average grades assigned to the event for subgroups of visitors divided on the basis of the duration of their presence at the event.

Age group	Grade	Number of respondents
19 years or younger	8.67	6
20-25 years old	7.77	31
26-30 years old	7.51	39
31-35 years old	7.55	31
36 years and older	7.31	13
<i>M</i>	7.63	120

Table 6. Average grades assigned to the event for subgroups of visitors divided on the basis of their age.

Interviews held with visitors during the event

The interviews held during Discovery07 included 124 visitors. 54% of this group was female, and 46% was male. About one third was in the age range between 26 and 30 years old, one quarter between 20 and 25, and another quarter between 31 and 35. The highest educational level was university for 71% of the group, high professional education for 23%, and middle professional education for 5% (and 1% other levels). 58% of the respondents had a job, 28% was studying, and 14% combined work and study.

On the first opinion question about what people thought about the event, 70% answered with positive judgments such as “interesting”, “fun”, and “very successful”; 11% of the respondents reported (in most cases at the beginning of the event) that the party “had not really started yet”, 8% thought that it was “reasonably okay”, and 4% found it somewhat “disorganized”.

Most interesting reported activities/elements of the event were the experiments (52%), speeches/presentations (16%), dance floor (9%), everything (7%) and bar (6%). Least interesting were: nothing (37%), experiments (20%), speeches/presentations (15%), interface (11%) and public/ambiance (6%).

With regard to the question about what respondents thought of the scientists present at the event 20% reported not to have seen a scientist yet and 28% not to have spoken to a scientist yet. 23% of the respondents perceived the scientists at the event as “sympathetic” and/or “nice”, 6% thought “they were just like ordinary people”, and 5% found them “too theoretical”.

The combination of science with art was judged positively by 76% of the respondents; these positive judgments were “interesting” (38%), “good” (31%), and “innovating” (7%). 11% of the respondents were negative about the combination of science with art.

The average judgment of visitors assigned by means of a (school-type) grade was 7.63. The average grade was somewhat lower at the beginning of the evening ($M = 7.52$ before 23.00 P.M.) than later ($M = 7.92$ after 01.00 A.M.), whereas the lowest average grade was given between 23.00 and 24.00 P.M. ($M = 7.10$). Grade level correlated most strongly with the duration of respondents’ presence: grades were higher as respondents were present longer (or grades were lower as people were present shorter) (see table 5).

Grades were related also with respondents’ age: as visitors were younger grades were higher (and as visitors were older, grades were lower) (see table 6).

Gender and educational level were not related to grade level.

As an illustration of what visitors said about the combination of science and art, the text box below presents some examples of statements of respondents.

- *Science and art go well together as they complement each other.
- *The combination is good for the development within both directions/cultures.
- *They work well together, as they may enrich each other. Science is very factual whereas art is creative. The combination stimulates the thinking about new prospects, providing a surplus value.

- *Especially music visuals (art) go well together with algorithms (science).
- *Art presents science at an understandable level.
- *Art is a way of transferring knowledge and science provides the up-to-date knowledge.
- *It's a good combination that sheds light on many issues and scientists are not mere nerds.
- *The concept is very cool; luckily scientists are not as dull as many people may think.

Conclusions and discussion

The present article was aimed at providing an example of how a public SC event can be evaluated. On the one hand, pre- and post-tests provided the possibility to investigate whether the event was effective. In this study the measurement of the effect pertained to the image of science and scientists. Of course, other evaluation studies may pertain to different types of possible effects such as changes in knowledge, attitude, and/or behavior. The advantage of using a pre-test and a post-test is that changes or effects can be established by investigating the differences between the two tests. In addition, because data were also collected among respondents who did not visit the SC event, there was the possibility of a comparison with a “control group”. On the other hand, the interviews held with visitors during the event gave the possibility of an evaluation in a more personal and qualitative manner. Solid effects of the event could not be measured with this method, but the “open” character of the interviews provided the possibility of getting a good idea of how visitors judged and perceived the event in a direct way. In addition, comparison of answers between subgroups of visitors helped to find out how different people responded to the event, and comparisons of reactions to different activities organized in the event could give an idea of how the public evaluated each of these activities.

The main goal of the present evaluation study was to investigate whether the image of science and scientists was improved by visiting the SC event. The results showed that the image was not improved. There were no significant differences between image scores established in pre- and post-tests, and this outcome was not affected when the dependent variables were corrected for exposure to other SC activities such as watching science programs on television, reading science articles in books or magazines, and visiting other science events. It may be concluded therefore that visiting the event did not lead to a more positive image of science and scientists. A possible explanation of the null-effect is that our sample was not very large. As the highest image scores for the 17- and 5-item image scales were found in the post-test among respondents who were present at the event, differences with the pre-test scores may have been statistically significant if the sample would have been much larger.

It was striking that the image of science scores in the present study were very high. First, they were much higher than those found among the public in an earlier Dutch study [1]. The high image scores found in the present study were more comparable with the scores found among science journalists in that study. These findings suggest that the visitors of Discovery07 deviated strongly from the general (Dutch) public in the respect that respondents in the present study had a much more positive attitude toward science and scientists than people in general. A probable cause for this positive image about science found in the present study may be that 79% of the participants in the web surveys and 71% of the visitors being interviewed had an academic education level, whereas of the general public in the Netherlands only about 15% has a university degree [2]. The finding that in the web survey of the present study images of science were more positive as the respondent's educational level was higher provides additional support for the validity of this explanation.

As reliable instruments were used to measure the image of science and scientists, the null-effect found in the present study cannot be ascribed to a possible low quality of the measurements. And although the

average image scores found here were very high, attribution of a null-effect to possible ceiling effects was also unlikely: for both measures the average item score was about 4 on a scale with a maximum of 5 leaving some space (although not much) to even higher average scores. A comment concerning the 17-item image of science scale developed by Krajcovich and Smith [4] is however that some of the statements included in this instrument seem to be based on somewhat stereotypical (and perhaps old-fashioned) ideas about science and scientists, which was at least apparent from some of the comments that participants in the web survey provided at the end of the survey. Future studies on the image of science should therefore perhaps consider the idea to develop a new or adapted version of this scale.

On the basis of the results of the interviews held with visitors of Discovery07 it may be concluded that most people highly appreciated the event. The average grade assigned to the event by the visitors was 7.63. Compared with the “usual grade” of 7 given by the Dutch public to, for example, television programs and public campaigns, and with much lower grades given to politicians, the grade given to the event was very high. The results of the present study showed that grades were higher when visitors were present longer at the event and when they had participated in more of the activities organized in the event. This finding may mean that the activities had a positive effect on the appreciation of the event, but could also be explained by the possibility that people who stayed longer at the event had already higher (pre-existing) opinions of science events or that higher opinions led to staying longer at the event. Two third of the sample of interviewed visitors were (very) positive about the way Discovery07 combined science with art, which makes it a combination to be considered for future use by other SC professionals at public science events.

The present study has of course its limitations. One limitation was that the sample in the survey study was small. Comparisons between pre- and post-test and in particular those between subgroups based on age and gender would have been more reliable and valid if the samples would have been larger. Another limitation was that the survey data were collected among groups of respondents that are perhaps difficult to compare. Although all of the respondents were in the same database of people who are (highly) interested in science and science events, the paying visitors who actually were present at the science event may have been more positive and motivated than the people who (eventually) did not show up at the event. In addition, because the respondents who participated in the pre-test were not the same respondents who participated in the post-test, the comparison of these groups may have been confounded by unknown variables. A design in which the same respondents would have participated in pre- as well as post-test would have been stronger, mainly because of the fact that in that case comparisons would have been more pure. Also, the expectation that a visit of a one-night SC event would lead to an effect on the public image of science may be characterized as naïve; changes in public attitudes toward science need much more time and effort. Although this may be true, the idea of the present study was to provide an example of how a SC event can be evaluated, including its limitations and merits. The type of evaluation conducted here may in the future be conducted for larger and longer-lasting SC activities. In addition, future research may profit from the knowledge of the limitations encountered in the present study. In that aspect, the present study hopefully contributes to future research in which the time and effort invested in SC activities will be evaluated in a reliable and solid way.

Acknowledgements

The author would like to thank the three interviewers: Christiaan Heijdenrijk, Frans van Konijnenburg, and Sietse Verhoog.

Notes and references

- [1] J.W. Becker and P.M. Van Rooijen, *Het beeld van de wetenschap in Nederland: Opvattingen van de bevolking, wetenschapsjournalisten, adviseurs, en Kamerleden in 2000*, [The image of science in the Netherlands: Opinions of the public, science journalists, advisors, and members of parliament in 2000], The Hague, the Netherlands, SCP/NWO (2001).
- [2] CBS, *Jaarboek onderwijs 2008* [Year book education 2008], The Hague, the Netherlands: CBS (2007).

- [3] Discovery07, *Internetsite voor het wetenschapsevenement* [Internet website for the science event] <http://www.discovery07.nl> consulted on September 26, 2007.
- [4] J.G. Krajcovich and J.K. Smith, *The development of the image of science and scientists scale*, *Journal of Research in Science Teaching* **19** (1982) 39-44.
- [5] C.E. Marshall, C.L. Blalock, Y. Liu, L.A. Pruski, M.A. Toepperwein, S.V. Owen and M.J. Lichtenstein, *Psychometric re-evaluation of the Image of Science and Scientists Scale (ISSS)*, *School Science and Mathematics* **107** (2007)149-160.

Author

Cees M. Koolstra is associate professor at the Department of Communication Science (Faculty of Social Sciences) at the VU University Amsterdam (the Netherlands). E-mail: cm.koolstra@fsw.vu.nl.