



UvA-DARE (Digital Academic Repository)

Post-observation conversations in the museum: using the self-evaluation of the supervisee as the starting point

Schep, M.; van Boxtel, C.; Noordegraaf, J.

DOI

[10.1080/09647775.2018.1501602](https://doi.org/10.1080/09647775.2018.1501602)

Publication date

2018

Document Version

Final published version

Published in

Museum Management and Curatorship

License

CC BY-NC-ND

[Link to publication](#)

Citation for published version (APA):

Schep, M., van Boxtel, C., & Noordegraaf, J. (2018). Post-observation conversations in the museum: using the self-evaluation of the supervisee as the starting point. *Museum Management and Curatorship*, 33(5), 506-523.
<https://doi.org/10.1080/09647775.2018.1501602>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)

Post-observation conversations in the museum: using the self-evaluation of the supervisee as the starting point

Mark Schep ^{a†}, Carla van Boxtel^b and Julia Noordegraaf^c

^aResearch Institute of Child Development and Education, University of Amsterdam, Amsterdam, The Netherlands; ^bResearch Institute of Child Development and Education, and Amsterdam School of Historical Studies, University of Amsterdam, Amsterdam, The Netherlands; ^cDepartment of Media Studies, University of Amsterdam, Amsterdam, The Netherlands

ABSTRACT

Post-observation conversations are often used to discuss the quality of teaching and education; however, little is known about the use of this method for the evaluation of teaching in museums. Drawing on interviews, in the first research question, we explored how museum guides ($n = 14$) and educators ($n = 8$) perceived a post-observation conversation when they used the guide's self-evaluation as the starting point. This is in contrast to the educator (observation) directed conversations that currently prevail in the participating museums, all located in the Netherlands. For the second research question, we coded seven conversations in order to investigate the guides' and educators' actual participation. Results indicate that guides felt a sense of ownership; furthermore, the participants evaluated the conversations as more equally balanced. Analysis of the conversations confirmed that there was a balanced interaction and that the guides greatly influenced the conversation's content.

ARTICLE HISTORY

Received 13 March 2018
Accepted 14 July 2018

KEYWORDS

Post-observation conversations; self-evaluation; museum guides; museum educators; observation; professional development


Introduction

Education has become a larger priority for museums in recent decades. As a consequence, these institutions are in the process of professionalizing their educational activities (Falk and Dierking 2013). Key to this transition are museum guides,¹ who engage in daily face-to-face interactions with visitors (Schep, Van Boxtel, and Noordegraaf 2018; Best 2012). Although studies on this topic are lacking, museums can stimulate the professional development of their guides and monitor the quality of their tours through observations, followed by post-observation conversations.

During a post-observation conversation, a supervisor and a supervisee discuss the performance of the supervisee after an observed performance. Such conversations (sometimes called meetings, conferences, mentoring conversations, or feedback conversations)

CONTACT Mark Schep  M.Schep@uva.nl; markschep@gmail.com  University of Amsterdam, Research Institute of Child Development and Education, PO BOX 15776, Amsterdam 1001 NG, The Netherlands

[†]Visiting address: Nieuwe Achtergracht 127 1018 WS Amsterdam.

 Supplemental data for this article can be accessed <http://dx.doi.org/10.1080/09647775.2018.1501602>.

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

are common practice in the field of teacher training and teacher education (e.g., Copland 2010; Ellis 1990; Kim and Silver 2016) and can serve different purposes. For example, they give feedback and help supervisees to further develop their skills; they can also help supervisors to come to an administrative decision.

This study is part of a larger project on the further professionalization of museum guides in three Dutch art and history museums located in Amsterdam: the Rijksmuseum (RMA), which focuses on Dutch art and history, from the Middle Ages onwards; the Van Gogh Museum (VGM), dedicated to the life and works of Vincent van Gogh; and the Stedelijk Museum (SMA), which focuses on modern and contemporary art and design. In this study, we focus on post-observation conversations in which the self-evaluation of the supervisee is used as a starting point. With this study, we want to contribute to the research on post-observation and the professionalization of museum personnel. The professionalization of guides is significant for guides and museums, but more important, it is relevant because guides are crucial to museum visitors' learning experience. This article continues with a discussion of the literature on post-observation conversations, followed by the research context of this study, and finally, we address the research questions.

Background

In recent decades, researchers have described several factors that influence the quality of a post-observation conversation. Sweeney (1983) identified four specific phases of the conversation. The process starts with careful (a) planning, but the conversation itself begins with (b) feedback that develops the (c) coaching and counseling session and concludes with (d) goal identification and closure. Sweeney emphasizes the importance of supervisors creating an atmosphere of trust in which they are able to provide feedback and to help the supervisee to identify areas in need of improvement as well as concrete goals.

In the planning phase, it is important that the supervisor considers the individual characteristics of the supervisee as well as his or her relationship with the supervisee. The available data for the conversation typically consists of the observations made by the supervisor, but data can also be derived from a video recording of the supervisee, the portfolio of the supervisee, feedback from students, or a self-evaluation. Supervisors must carefully examine the available data and make sure to focus on specific topics instead of trying to discuss every observed 'problem' during the conversation (Sweeney 1983).

Sweeney (1983) suggests beginning with an open-minded, purposeful exchange of ideas, in which the supervisee receives feedback but is given the opportunity to interact. During the conversation, the feedback must focus on specific events that occurred during the lesson as well as the overall approach of the supervisee. A difficulty for supervisors is managing the tension between assessment and development (Copland 2010). In recent decades, the focus in post-observation conversations has shifted from conversations with an evaluative goal to supervision that facilitates the personal development of the supervisee (Vasquez 2004). Supervisors should present ideas in the context of a helping relationship. In addition to managing people, it is important for supervisors to act as co-thinkers, inquirers, evaluators, and learning companions and to be a source for information (Orland-Barak 2006). It is the supervisor's responsibility to share his or her ideas about the performance of the supervisee and on how to improve; however, the supervisor must find a balance between being too directive and too nondirective (Sweeney 1983).

Due to the focus on personal development and collaborative inquiry, researchers stress the importance of supervisee involvement (e.g., Charlies et al. 2004; Tang and Chow 2007; Waring 2017), for example, through self-analysis (Ovando and Harris 1993). Self-reflection is essential for further learning and personal development (Shulman 1987; Veenman and Denessen 2001). Moreover, the insights gleaned from self-reflection appear to have a more positive effect on change of behavior than feedback from a superior (Deci and Ryan 2000). A recurring remark in the literature on post-observation conversations is the importance of dialogue to stimulate reflection during the conversation (e.g., Husu, Patrikainen, and Toom 2007; Walsh 2011, 2013).

There are some factors that can negatively influence the quality of the conversation, including supervisee resistance (Waite 1995), the asymmetrical role relationship, supervisee anxiety (Vasquez 2004), and supervisor dominance during the conversation (Hyland and Lo 2006). Therefore, it is important to manage the expectations of the supervisee. Chamberlin (2000) noted that when supervisees expect positive appraisal and constructive criticism, they can feel disappointment when the conversation leans more towards reflection rather than evaluation. Most of these pitfalls can be prevented by the success factors described in the literature, such as careful planning, mutual trust, clarity about the goals, a supervisor's good communication and interpersonal skills, and when the supervisee adopts an active role (Charlies et al. 2004; Kim and Silver 2016; Ovando and Harris 1993; Sweeney 1983).

Context of the study

The three partnering museums in this research project organize short training programs lasting 2–5 days that prepare guides for their educational task. In addition, the museums offer workshops and lectures and have annual post-observation conversations to monitor the quality of their guides. In these museums, the educator typically makes notes during the tour and has a short list of focus points (i.e., the structure of the tour, goals, competencies), which are more or less chronologically ordered. The post-observation conversation that follows is often conducted immediately after the observed tour and takes approximately 30–60 min. In most cases, the observations of the educator steer the direction of the conversations. Educators in our study indicated that there is a correlation between the quality of the tour and the length of the meeting – the better the tour, the shorter the conversation. Throughout a conversation, educators will emphasize prominent situations that occurred during the tour, mainly the things that went wrong or that might be improved. At the end of a conversation, goals for improvement are identified and discussed. For instance, when a guide's performance is deemed insufficient, the educator will schedule a follow-up observation and subsequent conversation.

Possible improvements to current practices

Based on the above-described literature and the observed practices in participating museums, we will discuss possible improvements to post-observation conversations. Grenier (2009) examined in what way (unpaid) guides in history museums develop expertise. She concluded 'learners need to take charge of their own development through continuing education, self-directed learning, and reflection on prior experiences in order to

move from competence to expertise' (154–155). The capacity to reflect and to learn from experiences is crucial for professionals (Marienau 1999); therefore, museums could improve post-observation conversations by involving museum guides in their own professional development.

Researchers suggest the use of standards/rubrics as a common language for talking about teaching (Ross and Bruce 2007; Sergiovani and Starratt 2002). There are several advantages when the conversation is facilitated by such written documents; for example, it can help to recall situations, it can serve as a catalyst for reflection and discussion, and it can structure the conversation (Engin 2015; Katic, Hmelo-Silver, and Weber 2009).

Along the same lines, a self-evaluation can function as a common language; it also ensures that guides practice self-reflection. In a former study (Schep, Van Boxtel, and Noordegraaf 2018) validated by museum education professionals, researchers created a competency profile for museum guides, which makes it possible to develop a self-evaluation instrument. In the current study, this self-evaluation served as the starting point of the post-observation conversations. Further specifications of our self-evaluation instrument are discussed below. Our aim was to evaluate the use of this tool as the starting point for a post-observation conversation, which resulted in the following research questions:

1. How do museum guides and educators perceive a post-observation conversation based upon the self-evaluation of the guide?
2. How does the museum guide participate in the conversation when his or her self-evaluation is used as the starting point?

We expected that the use of a self-evaluation would give museum guides a more active role in these post-observation conversations and thus in their own learning processes.

Method

To study the use of the self-evaluation instrument in post-observation conversations we used a mixed-method design. In order to answer the first research question, we interviewed all participants. Then we used the video recordings of the post-observation conversations to analyze the actual contribution of the museum guides in the conversation. We first tested the whole procedure during a pilot round.

Setting and participants

All participants were affiliated with the three partnering museums in this research project – the RMA, SMA, and VGM. These museums are based in Amsterdam and are some of the largest museums in the Netherlands.² Thus, they have specialists for different educational tasks, such as the supervision of their guides. In smaller museums, the staff often consists of one or two educators, who have a broader range of responsibilities. The partnering museums each have between 30 and 110 guides who frequently give tours. Guides are hired as freelancers and most of them work in at least two museums. We deliberately selected paid professionals because we consider tour guiding a profession.

A total of eight museum educators, six women and two men, participated in this study. Three educators participated in both the pilot study and the experimental study. Their experience varied from 1 to 10 years ($M = 6$ years), and their ages ranged from 27 to 59 years ($M = 41$ years). All educators had, at minimum, a bachelor's degree. At the SMA, one educator was responsible for all the observations and conversations. The RMA and the VGM divided the task between four and three educators, respectively. These educators were asked to nominate museum guides in three categories of experience: less than 3 years, 3–8 years, and over 8 years. For the pilot study, we selected three experienced museum guides, all women, who were expected to be capable of giving feedback on the tools and on the procedure.

Initially, we sent a request for participation to 15 museum guides. Whenever a guide declined to participate, we approached another guide from the list until 15 guides (five in each category of experience) agreed to participate. One guide was not able to complete the full procedure because the educator transferred to another museum; therefore, 14 museum guides participated in total. The experience of the guides, 12 women and two men, varied from four months to 12 years ($M = 5.5$ years), and their ages ranged from 26 to 58 ($M = 39$ years). All guides had, at minimum, a bachelor's degree. Participants signed an informed consent form and had the opportunity to quit the study at any time.

Instrument development³

Self-evaluation tool

Self-evaluation is an internal process of self-reflection and self-assessment. In their comparative analysis of the two concepts, Desjarlais and Smith (2011) define reflection as the 'personal process that deepens one's understanding of self and can lead to significant discoveries or insights' (p. 3) and self-assessment as 'the process that involves establishing strengths, improvements and insights based on predetermined performance criteria' (p. 3). Various methods are used for self-evaluation, such as questionnaires, rating systems, daily journals, and portfolios (Struyk and McCoy 1993). The design of our self-evaluation tool was based on the 45 competencies for museum guides in art museums and history museums, which we defined in a previous study of this project (Schep, Van Boxtel, and Noordegraaf 2018). The self-evaluation consists of 45 competencies, organized into four areas of competence: (a) handling the group within the museum environment, (b) communication skills, (c) knowledge and pedagogy, and (d) professionalism. The museum guides must rate themselves on a five-point Likert scale ranging from 1 (*insufficient*) to 5 (*excellent*). For each of the four main areas of competence the guides are asked to identify one or two strengths and aspects needing improvement. No important changes to the instrument were made after the pilot study in which we interviewed three educators and three museum guides.

Observation tool

The observation tool contains the same 45 competencies as the self-evaluation and is organized into the four areas of competence. The competencies are only described by their short title and are formulated as observable acts. For example, we changed *enthusiastic and energetic* in the self-evaluation to *is enthusiastic and energetic* in the observation tool. The tool contains space for notes and the same five-point Likert scale to help rate

the competencies of the guide. Just like the guides, the educators were asked to indicate strengths and aspects needing improvement for each area. The pilot study did not result in changes. Points for discussion were the anchors of the scale, which ranged from *insufficient* to *excellent*. One educator opted for different terminology because she felt uneasy with the terms; we did not address her suggestion because the scale was generally shared among the other participants.

Conversation guideline

In order to structure the conversation between the museum educator and the museum guide we provided guidelines based on the four areas of competence. The guides started the conversation by naming their strengths and aspects needing improvement in the first area of competence. Next, the educator had the opportunity to respond to these topics. When both participants felt the first area was discussed sufficiently, the second area of competence was discussed in the same way, and so on. In the last phase, we instructed the participants to set goals.

Procedure

First, the guides filled in the self-evaluation—based on their general performance—and sent it to the researcher as well as the educator of their museum. Next, the museum educators used the observation tool to observe two tours: one for primary education students and one for secondary education students. Therefore, the observations—contrasting the self-evaluation—were based on two specific tours. Subsequently, the museum guide and museum educator engaged in a post-observation conversation, scheduled according to the convenience of the participants. The educator brought a hardcopy of the self-evaluation to the conversation without reading it beforehand.

Data gathering and analysis

Interviews

In order to answer the first research question, ‘how do museum guides and educators perceive a post-observation conversation based upon the self-evaluation of the guide,’ we interviewed all 25 participants using a semistructured interview format. We asked participants about their experiences with the instruments, the use and usability of the instruments during the conversations, how they perceived the conversation using the self-evaluation as the starting point, and to compare this with conversations in which the observations of the educator were the main input. We recorded, transcribed, and analyzed the interviews using Atlas.ti. The coding consisted of two phases: first, we grouped the answers of the participants into the categories based on the interview questions; next, we coded major themes that emerged within the categories in the second round (Hammersley and Atkinson 1995). This resulted in the main themes: organization, complete picture of performance, agency, ownership, balanced feedback, equally balanced conversation, structure of conversation, common language, and objective standard.

Post-observation conversations

All conversations lasted 45–75 min; they were recorded on video, transcribed, and coded using Atlas.ti. At the SMA, one educator was responsible for all four conversations; at the

RMA, four educators engaged in a total of five conversations; and at the VGM, one educator had three conversations and two others engaged in one conversation. From seven different educators, we obtained at least one recording of a post-observation conversation.⁴ In order to ensure that the results were not overly influenced by educators who were engaged in more conversations, we only selected one educator from each conversation for further analysis. When more recordings of an educator were available, we selected the one in which the prescribed format was most clearly used.

We coded the conversations in three steps. First, we coded strengths and aspects needing improvement introduced by the participants. This coding addressed to what extent both participants initiated topics in the conversation. Second, these strengths and aspects needing improvement were coded for competency title (45 competencies). Third, for each introduced topic (strength or aspect needing improvement), we coded whether the topic was further discussed or not. This coding provided more information about the extent to which the museum guides influenced the content and the conversation. The cut-off to decide if an initiated topic was discussed was when there were at least two exchanges about the specific topic. For an overview of the different codes and an example of the coding see [Tables 1](#) and [2](#), respectively.

An additional way to explore the contribution of the participants is to count their words. Vasquez (2004) shows that in typical post-observation conversations the proportion of talk by the supervisor (approximately 80%) is much more than that of the supervisee. In Vasquez's study, six conversations were analyzed; all supervisees were teaching assistants. The proportion of talk is based on the total amount of words used by each participant.

In order to investigate the inter-rater reliability, the second author coded three conversations. Both researchers identified segments in which the educator or the museum guide

Table 1. Coding scheme.

Main category	Subcategories	Code
1. Introduction of topics	Museum educator introduces strength	ME-STR
	Museum guide introduces strength	MG-STR
	Museum educator introduces aspect needing improvement	ME-IMP
	Museum guide introduces aspect needing improvement	MG-IMP
2. Specific competency ^a that is introduced	Competency open attitude	COMP-ATT
	Competency making an immediate rapport	COMP-RAP
	Competency assessing the group	COMP-ASS
3. Discussion of introduced topics	Discussed	DISC
	Not discussed	N-DISC

^aFor a complete list of competencies see supplemental materials.

Table 2. Example of coded transcriptions: the museum guide introduces an aspect needing improvement (competency flexibility) which is discussed.

Educator or Guide	Utterance
Guide	With the adults I thought, we can't be at the same object with two groups, but I don't have many other options yet. That's something I have to learn. So I think, okay, we go a bit later, that's not a problem. It's a point for improvement, to know more, because then I can be more flexible. I'm not really flexible yet.
Educator	That's what I noticed as well. That has to do with getting more hours under your belt.
Guide	I only gave seven tours so far.
Educator	Exactly, you have to fill your backpack with objects and other routes. But that comes with time.

introduced a topic; in total, 108 segments were identified. The percentage of agreement over these segments was 75%. In the three conversations, raters agreed on 61 instances in which the guide initiated a topic and 20 instances in which the educator initiated a topic. Consequently, both coders coded whether or not these initiated topics ($N = 81$) were discussed. The observed agreement was 90% (Cohen's Kappa = .80), which is considered substantial (Landis and Koch 1977) and excellent by Fleiss (1981). Inconsistencies were mostly because one of the raters failed to identify an introduced topic, coded a segment with a different competency, or assigned the code to the wrong participant.

Results

The Results section consists of two parts. First, we describe the perceptions of the participants in order to answer the first research question, 'how do museum guides and educators perceive a post-observation conversation based upon the self-evaluation of the guide?' The above-mentioned main themes—organization, complete picture of performance, agency, ownership, balanced feedback, equally balanced conversation, structure of conversation, common language, and objective standard—are italicized. Second, we present the content analysis of the conversations in order to answer the second research question, 'how does the museum guide participate in the conversation when his or her self-evaluation is used as the starting point?'

Perceptions of the participants

Using the observations as the starting point

We asked all participants to describe how post-observation conversations usually take place and to evaluate the value of these conversations. Overall, most museum guides evaluated the post-observation conversations positively – mainly because they acknowledged that it is very useful to get feedback on their performance. These conversations are one of the rare opportunities to get (professional) feedback. However, museum guides also explained that some conversations are one-sided. This causes stress because the guides feel that they are being assessed and judged and because they often do not know what to expect. For example, Amber, a guide who participated in the pilot study, explained:

I think that there was more space for my voice. Normally it is just more one-way communication: the educator tells you what they like ... the thing I find annoying in such conversations is that I always have to be careful not to get defensive.

Educator Sophie added, 'In the way we usually work it was still very much that I told them [guides] what I saw and they could respond on that. Next, I advised them what options they have.'

Using the self-evaluation as the starting point

Most participants evaluated the use of the self-evaluation in the conversations positively. Only two organizational points were mentioned as possible setbacks: the self-evaluation consisted of multiple pages, which can be inconvenient because it precludes a quick overview; and some participants indicated that the whole procedure was more time

consuming. A main strength of the new format was that all participants thought that the conversation was a more accurate indicator, which provided a complete picture, and was not based on only one particular tour. For example, educator Sophie stated:

Firstly, what I think is very good here, especially when using the self-evaluation as a starting point and then responding based on your observations, is the fact that it ensures that the conversation goes beyond the particular tour you observed, and it really concerns the general performance and development of the guide. In this way, the observations of that particular tour are illustrative or serve as supplementary materials.

Museum guide Kate, who participated in the pilot study, made a similar point, by stating her objections to the usual post-observation conversations: 'It's not about how I give tours. It's about that specific strange thing that happened during the last tour, and that's a pity, but, of course the observations during this particular tour are the only tool the educator has.'

Another strength was the sense of *agency* the guides experienced. For example, guide Natalie stated, 'It [the self-evaluation] gives structure to the conversation, you are not constantly defending yourself, but it is more objective about certain topics; that's what I liked, and you get the opportunity to give your opinion about a situation.' Guides explained that because the self-evaluation was used as input for the conversation, they were stimulated to reflect on their own practices, which prepared them for the conversation and also led to a *more equally balanced conversation*. Guide Michael expressed this as follows, 'Normally, it feels more like you are being judged, but because in this situation you already reflected on your own practices, you are sort of a step ahead, which makes the conversation more equal.'

Museum guide Rose argued that guides are more eager to work on their professional development when it is based on their own reflection. Guides explained that the self-evaluation helped them to find the aspects they needed to improve and that the educator could help them find solutions during the conversation, which gave the guides more *ownership*. Educator Iris stated something similar:

This format worked very well because the guide reflected on his own competencies and came up with points that I also observed. I think it's much stronger when a guide comes up with points for improvement stemming from self-reflection than when someone from the office confronts him with these points.

For museum guide Beth, however, the self-evaluation did not add any value to the conversation. She stated that the most important attribute for an educator was that he or she be an expert and trustworthy. She explained that she would invest the time spent on the self-evaluation differently, for example, by shadowing other museum guides. However, Beth liked it when she identified two competencies as possible aspects for self-improvement, while the educator's assessment indicated that she is competent in these. This was the case for other guides, who felt insecure about a certain competency but were, in some cases, reassured by the educator.

Related to the issue of trust, it is important to decide which topic to discuss in a conversation. Educator Eliza argued that she normally must choose which areas of improvement they want to address. The conversation will not result in a change of behavior if the museum guide is overwhelmed by criticism or advice. In this instance, the use of the self-evaluation helps educators and guides to find a *balance* between the qualities

of the guides and the competencies that need improvement. Eliza further posited that the self-evaluation might especially help less experienced or unskilled supervisors because it provides structure to the conversation.

Key to the conversation was the list of competencies. Both the educator and the guide were familiar with the competencies; they felt that the list served as a clear and *objective standard* in the conversation and helped them to speak a *common language*. For example, educator David explained, ‘the fact that you speak the same language, and that you are transparent on what you assess, and thus on which points the guides can develop ... in this, ownership of the guides is very important.’ Participants explained that using the list of competencies and the self-evaluation clarified that the conversations are less about the interpretation or personal preferences of an individual observer. Although interpretations about particular behavior still can differ, using the list of competencies reduced misunderstandings, according to the participants.

Analysis of participants’ contribution to the conversation

We already mentioned that the participants perceived that using the self-evaluation instrument as a start for the conversation contributed to an equally balanced conversation. In this section we present the results of our analysis, which confirms the picture painted by the participants.

The proportion of talk during the seven conversations was almost equally divided between the two participants; on average, the guides spoke 50.2% of the time. See [Appendix A](#) for an overview. In four cases, the guide spoke collectively longer than the educator. In the least balanced conversation between guide Lucia and educator Claire, Lucia spoke 37% of the time, which still is more than the average of 22% for supervisees in a study by Vasquez (2004).

The data on the introduced topics during the conversation, as presented in [Table 3](#), show that the museum guides introduced more topics than the educators: 61% of the topics were introduced by guides. Additionally, [Table 3](#) shows that in the seven conversations 62% of the discussed topics were introduced by a museum guide and that in five of the seven conversations most of the discussed topics were introduced by the guide. Only in Conversation 5 did the museum educator clearly dominate the content of the discussed topics. In this conversation, the educator initiated 75% of the discussed topics. The differences between the conversations are discussed below.

Table 3. Frequencies of topics introduced and topics that are discussed and asymmetry.

Conversation	Topics introduced by		Asymmetry %-%	Discussed topics introduced by		Asymmetry %-%
	Guide	Educator		Guide	Educator	
1. Emily and Anna	33	7	82–18	14	5	74–26
2. John and Iris	25	14	64–36	16	7	70–30
3. Lucy and Sophie	16	10	62–38	14	6	70–30
4. Rose and Joanna	17	14	55–45	5	3	62–38
5. Lucia and Claire	10	21	32–68	4	12	25–75
6. Martha and Eliza	28	7	80–20	15	3	83–17
7. Mary and David	9	14	39–61	6	10	38–62
Total	138	87	61–39	74	46	62–38

Note: The first name is the name of the guide; the second name is the name of the educator.

We found that the majority of the subsequent topics discussed were aspects needing improvement (61%); strengths introduced by educators were barely discussed (26%). It must be noted, however, that some of strengths introduced by guides were discussed as an aspect needing improvement. The best example is the competency 'creating a common thread', which was a point for discussion in many conversations.

An example of how the self-evaluation played a role in the conversation and how the guide influenced the conversation is given in Excerpt 1. Based upon the self-evaluation ('points for improvement'), the museum guide introduced a topic using the language that is used in the list of competencies. In this list, the competency 'positioning' is described as

the guide is flexible and capable of dealing with the busyness in the museum. The guide positions him or herself and the group in a way such that all the students can see the object without blocking the passage for the other visitors.

In the excerpt below, contrary to the self-evaluation of the guide, the educator explains that she observed that the guide demonstrated *flexibility* and the capacity to deviate from his route.

Excerpt 1

John: I find it difficult to deviate from my route. [...] Points for improvement: *positioning* and deviate from the route. I think this is because of a lack of experience, at least in this museum. But you will learn that in the process. It improves every time. Especially when it is very busy, although you really want to go there. When you have more knowledge it is easier to deviate from your route. Then you just have some other objects in the same category. That gives me the feeling I have to improvise, and I think that decreases the quality.

Iris: [...] What I find funny, is that you express that you have difficulties deviating from the route. What I observed in the first tour, what was striking, that you asked, 'have you been to this museum before?' Some of them did. Then you asked the girl, 'where do you want to go?' That really felt like you were improvising on the spot, but maybe that was already in your routing?

John: It really depends on the group.

Iris: It was a beautiful moment, because the students really influenced what they were about to see.

Excerpt 2 includes an example in which educator Anna introduces a point for improvement, 'asking questions,' and gives examples about these points as well as advice. The guide Emily, a novice at this museum, was quite passive in this part of the conversation but demonstrated a willingness to learn. The educator and guide spoke about competence Area 3, 'knowledge and pedagogy,' and because the guide did not introduce this competency, the educator decided to do so. This excerpt illustrates the typical conversations involving novice museum guides.

Excerpt 2

Anna: Sometimes I noticed some fake interaction. At one object you had a roll of paper and asked, 'what is inside?' Probably they don't have a clue. Then you said, 'it starts with a P.' That does not really add something, a guessing game. It's not relevant.

Emily: Okay.

Anna: You also ask for facts, things they can't know, which is also not relevant. For example, 'who is the architect of this building?'

- Emily: Ah, yes.
- Anna: 'Binder, what sort of binder is there?' 'Who painted something?' 'In what year was the VOC founded?'
- Emily: Isn't that something the older ones should know?
- Anna: No, if you ask a random adult, they also don't know.
- Emily: I should write this down. [...]
- Anna: They really don't know. [...] So, about the questions, more open questions, stimulate them to look, ask questions with more than one answer, explorative questions.

The way in which the instruments were used and whether guidelines were followed influenced many conversations in our study. For example, in two conversations (Emily and Anna; Martha and Eliza), instead of naming only one or two strengths per category, as the guidelines suggest, the guides more or less touched upon all the competencies. The number of strengths addressed by the experienced Martha and the less experienced Emily was more than average – 22 and 23, respectively, compared to other guides' average of eight. Emily, a novice guide in the museum, mostly made statements such as, 'cooperating with colleagues, yes, I think so. Cooperation with security as well, if they say something 'I do it,' cooperating with teachers, I think that's good as well.' We coded these fragments as introduced topics, which significantly increased the number of topics introduced by Emily (and included a lot of coded strengths). In the conversation between Emily and Anna, 19 introduced topics were discussed further, 12 of which concerned aspects needing improvement. Furthermore, some of the strengths introduced by guide Emily were discussed as aspects needing improvement.

During Conversation 4, Rose and Joanna discussed only eight topics, which is relatively few compared to the other conversations that averaged 17 discussed topics. Rose, an experience guide, discussed aspects needing improvement in depth; this reflected the prescribed format, which suggests discussing only the main strengths and aspects needing improvement. Compared to conversations of other experienced guides (Conversations 3, 6, and 7), Rose and Joanne neglected to broadly discuss strengths, even though both participants introduced them.

The experience of the guide possibly influenced the conversation. Three conversations (Conversations 1, 2, and 5) involved museum guides new to a particular museum, though John had 4 years of experience in other museums. Compared to the experienced guides' conversations, participants in Conversations 1, 2, and 5 more often explored aspects needing improvement (an average of 14 compared to eight). In guide Lucia's case, educator Claire mainly introduced the aspects needing improvement and often gave advice on specific museum guidelines and on what to do or what not to do in a particular situation (see Excerpt 3). Mostly, these were minor points that will likely improve in time as Lucia gains experience.

Excerpt 3

- Claire: A minor thing. I have some tips for the first tour. Also, because you are new.
- Lucia: Yes, please.
- Claire: Immediately, when you write the children's names on the sticker, you should say that they have to stick it on their shirt, and don't touch it anymore. I saw them exchanging their name tags.

In some cases, it was difficult to evaluate who introduced a topic. Segments encoded as 'introduced by the guide' were those introduced verbally by the guide, but also included those introduced by the educator if he or she was reading from the guide's evaluation sheet. During the conversation between educator David—who had less than 1 year of experience as an educator, and until recently was a guide himself – and guide Mary, it was David who used the self-evaluation to start the conversation. In David's three other conversations, we saw the same pattern. Nevertheless, the guides David interviewed indicated that the conversations were equally balanced. This is supported by our analysis of the conversation between Mary and David. [Table 3](#) shows that six of the nine topics introduced by Mary were discussed further, which suggests that the guide influenced the content of the conversation.

Discussion

The purpose of our study was to gain insights into the use of a self-evaluation as the starting point for post-observation conversations between museum guides and museum educators. The first research question focused on participants' perceptions about the post-observation conversations based upon the self-evaluation of the guide. Participants positively evaluated the provided tools and conversation guidelines. Some of the characteristics of successful conversations and possible pitfalls mentioned in the literature can be identified in this study. As a result of the instruments used, the main advantages of our format included the perceived ownership by the guides, the perceived equality between the participants of the conversation (e.g., Charlies et al. [2004](#); Kim and Silver [2016](#)), the sense of objectivity, and the use of a common language (e.g., Sergiovani and Starratt [2002](#)).

The second research question addressed the actual contribution of the participants. Analysis of video recordings showed that the guides were involved in the selection of the topics that were discussed. Furthermore, guides indicated that the self-evaluation made them reflect on their own practices, which had two advantages according to the participants: (a) they were prepared for the conversation and (b) some of the guides and educators acknowledged that insights stemming from self-reflection probably had more impact on guides' willingness to engage in further action, which supports the findings of Deci and Ryan ([2000](#)).

The format also helped to structure the conversation (e.g., Engin [2015](#); Katic, Hmelo-Silver, and Weber [2009](#)). An educator explained that it was not her job to be judgmental. She was in agreement with colleagues that educators should act as coaches, be supportive, and be open to the views of the guide. This is consistent with the literature on the role of supervisors in post-observation conversations (e.g., Orland-Barak [2006](#); Sweeney [1983](#)).

One of the limitations of this study was that we did not collect empirical data on post-observation conversations that did not use a self-evaluation instrument. However, the participants described how conversations usually occur and compared them to conversations that used our format. The educators and guides painted a consistent picture about the pros and cons of conversations that do not use a self-evaluation instrument, which suggests that we had a clear image of the situation before our study.

Another possible confounding factor is that our guidelines were not used in all conversations as prescribed. In some cases, almost all 45 competencies were addressed, but the

participants should have focused on the strengths and one or two aspects needing improvement in each main area. In one case, the educator occasionally used the guide's self-evaluation sheet to start the conversation. Furthermore, we had to decide upon a boundary for coding segments as 'discussed'. We did not make any difference between topics that were discussed elaborately, or topics that included only a couple of turns. Future studies could further explore the 'quality' and 'quantity' of exchanges during post-observation conversations. Nevertheless, based on the interviews and on the analysis of the conversations, it is reasonable to conclude that the conversations were equally balanced and that there was more space for the voice of the guide – compared to conversations directed by the observations of the educator. When giving the instructions to participants, it is important to emphasize that the discussion should focus on the main strengths and aspects needing improvement.

We tested the tools on professional museum guides in three of the largest museums in the Netherlands. Both the novice and the experienced guides evaluated the self-evaluation as a valuable instrument to aid professionalization. In a former study (Schep, Van Boxtel, and Noordegraaf 2018), the same competencies were validated by a panel of experts from museums across the Netherlands. It may be that observations and post-observation conversations are less common in museums that work with volunteers. Nevertheless, these volunteers can use the self-evaluation to indicate their areas needing improvement and to ask their supervisors for feedback and advice.

Professionalization is important for museum guides and for museums. More important, it is relevant because guides are a crucial factor in museum visitors' learning experience. The findings from this study can contribute to the professionalization of guides, the result of which should help to improve museum tours. The instruments were developed especially for the museum environment; a teaching environment that has its own particularities. In two former studies (Schep, Van Boxtel, and Noordegraaf 2015; Schep, Van Boxtel, and Noordegraaf 2018), we focused on the museum as a learning environment, and the specific competencies that guides need. For example, school groups often only visit once; (authentic) objects are central; tours last only 60–90 min; other visitors wander around; and the goals of the tours balance between fun and educational. This demands a wide range of competencies, such as flexibility, the capability of making immediate rapport, knowledge about the target group, communication and pedagogical skills, and content knowledge. Furthermore, the guides in this study are freelancers; this in contrast to school teachers who often have a permanent contract. Therefore, for the guides there is more at stake. All these characteristics affect the evaluation and the process of a post-observation conversation. The fact that the self-evaluation was used as the starting point of the conversation helped to balance the playing field, according to the participants.

The results of this study can be used in numerous ways. The complete procedure can help educators to monitor guides and to support guides in their professional development. Moreover, guides can use the self-evaluation to reflect on their own practices. Museums can also use the instrument to get an overview of guides' competencies, which can help museums pair novice guides to more experienced guides – this informal mentorship arrangement was suggested by museum guide Rose. For example, guides who need to improve their pedagogical skills can be advised to observe guides who excel in these competencies. Consistent with Grenier (2009), most guides in our study agreed that learning by shadowing other guides is a valuable way to acquire new insights.

In our study, the educators did not read the guides' self-evaluations beforehand because we wanted an open conversation in which the insights of the guides formed the starting point. This can be done differently; for example, the educators could read the self-evaluation before observing the tour and focus on the competencies that are indicated by the guide. In this way, the educator has the opportunity to think about feedback on these topics.

It must be noted that the tools and the format were not developed to function as an assessment tool. The instruments can help museum guides reflect on their own practices, can help them to identify their own strengths and aspects needing improvement, and can stimulate their professional development. We advise each museum to adjust the tools to their own preferences, for example, to select the competencies that are most relevant to them.

This study expands the current research on post-observation conversations and the research on the professional development of museum guides – particularly on the use of self-evaluation during post-observation conversations. Post-observation conversations guided by the self-evaluation of supervisees contribute to a more structured conversation and can help supervisees control the content of the conversation, which may lead to a more valuable learning experience.

Notes

1. In this study, we use the term *museum guide*, which refers to persons who give tours to visitors and thus are engaged in daily face-to-face interactions with visitors. In our study, the guides are paid freelancers and not volunteers. Other frequently used terms for this profession are, for example, facilitator, tour guide, museum docent, or museum educator. When we speak of *museum educators* we mean paid members of the staff who are responsible for the development of the educational programs and the supervision the museum guides.
2. All instruments and interviews were in Dutch, therefore the presented materials and quotes are translations.
3. All instruments are added as supplementary materials and are available via <https://figshare.com>.
4. One of the conversations was not properly recorded and could not be used.

Acknowledgement

The authors would like to thank the funders—Fonds 21 and the Mondriaan Fonds—the partnering museums, and most of all, the museum educators and museum guides who participated in this project.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Mark Schep has a Master of Science in Social Psychology and Master of Arts in History Education at the University of Utrecht. Currently he is a Ph.D. candidate at the University of Amsterdam on the research project *Professionalization of museum guides and improvement of school tours in art and history museums*. In this project Schep examined which competencies museum guides should

have when they give guided tours to school groups, and if and how observation instruments and self-evaluation instruments can help the further professionalization of museum tour guiding as a profession.

Carla van Boxel is professor of History Education at the Research Institute of Child Development and Education and the Amsterdam School for Historical Studies, University of Amsterdam. She gives – together with Julia Noordegraaf – lead to the research project *Professionalization of museum guides and improvement of school tours in art and history museums*. Her main research interest is the learning and teaching of history in schools and museums, with a focus on the improvement of students' historical reasoning. She published in international volumes and peer reviewed journals, such as *Cognition and Instruction*, *Journal of Curriculum Studies*, and *Educational Psychology Review*.

Julia Noordegraaf is professor of Digital Heritage and director of the Amsterdam Centre for Cultural Heritage and Identity at the University of Amsterdam. Her research focuses on the preservation and reuse of audiovisual and cultural heritage. She has published, amongst others, the monograph *Strategies of Display* (2004/2012) and, as principal editor, *Preserving and Exhibiting Media Art* (2013). Noordegraaf currently leads research projects on the conservation of digital art and on the reuse of digital heritage in data-driven historical research. She acts as board member for Media Studies in CLARIAH, the national infrastructure for digital humanities research, funded by the Netherlands Organization for Scientific Research, NWO.

ORCID

Mark Schep  <http://orcid.org/0000-0002-8713-4452>

References

- Best, K. 2012. "Making Museum Tours Better: Understanding What a Guided Tour Really is and What a Tour Guide Really Does." *Museum Management and Curatorship* 27 (1): 35–52. doi:10.1080/09647775.2012.644695.
- Chamberlin, C. T. 2000. "TESL Degree Candidates' Perceptions of Trust in Supervisors." *TESOL Quarterly* 34 (4): 653–672.
- Charlies, S., L. Ria, S. Bertone, J. Trohel, and M. Durand. 2004. "Interactions Between Preservice and Cooperating Teachers and Knowledge Construction During Post-Lesson Interviews." *Teaching and Teacher Education* 20 (8): 765–781. doi:10.1016/j.tate.2004.09.001.
- Copland, F. 2010. "Causes of Tension in Post-Observation Feedback in Pre-Service Teacher Training: An Alternative View." *Teaching and Teacher Education* 26: 466–472. doi:10.1016/j.tate.2009.06.001.
- Deci, E., and R. Ryan. 2000. "The 'What' and 'Why' of Goal Pursuits: Human Needs and the Self Determination of Behavior." *Psychological Inquiry* 11 (4): 227–268. doi:10.1207/S15327965PLI1104_01.
- Desjarlais, M., and P. Smith. 2011. "A Comparative Analysis of Reflection and Self- Assessment." *International Journal of Process Education* 3 (1): 3–18. <http://www.ijpe.online/2011/reflection.pdf>.
- Ellis, R. 1990. "Activities and Procedures for Teacher Preparation." In *Second Language Teacher Education*, edited by J. Richards, and D. Nunan, 26–36. Cambridge: Cambridge University Press.
- Engin, M. 2015. "Written Artefacts in Post-Conference Feedback Sessions: The Running Commentary as a Support for Teacher Learning." *Journal of Education for Teaching* 41 (3): 254–266. doi:10.1080/02607476.2015.1041289.
- Falk, J., and L. Dierking. 2013. *The Museum Experience Revisited*. Walnut Creek, CA: Left Coast Press.
- Fleiss, J. L. 1981. *Statistical Methods for Rates and Proportions*. New York: John Wiley.
- Grenier, R. S. 2009. "The Role of Learning in the Development of Expertise in Museum Docents." *Adult Education Quarterly* 59 (2): 142–157. doi:10.1177/0741713608327369.
- Hammersley, M., and P. Atkinson. 1995. *Ethnography: Principles in Practice*. London: Routledge.

- Husu, J., S. Patrikainen, and A. Toom. 2007. "Developing Teachers' Competencies in Reflecting on Teaching." In *Making a Difference: Challenges for Teachers, Teaching and Teacher Education*, edited by J. Butcher, and L. McDonald, 127–140. Amsterdam: Sense.
- Hyland, F., and M. Lo. 2006. "Examining Interaction in the Teaching Practicum: Issues of Language, Power, and Control." *Mentoring & Tutoring: Partnership in Learning* 14 (2): 163–186. doi:10.1080/13611260500493535.
- Katic, E. K., C. E. Hmelo-Silver, and K. H. Weber. 2009. "Material Mediation: Tools and Representations Supporting Collaborative Problem-Solving Discourse." *International Journal of Teaching and Learning in Higher Education* 21: 13–24.
- Kim, Y., and R. E. Silver. 2016. "Provoking Reflective Thinking in Post Observation Conversations." *Journal of Teacher Education* 67 (3): 203–219. doi:10.1177/0022487116637120.
- Landis, J. R., and G. G. Koch. 1977. "The Measurement of Observer Agreement for Categorical Data." *Biometrics* 33 (1): 159–174.
- Marienu, C. 1999. "Self-Assessment at Work: Outcomes of Adult Learners' Reflections on Practice." *Adult Education Quarterly* 49 (3): 135–146. doi:10.1177/074171369904900301.
- Orland-Barak, L. 2006. "Convergent, Divergent, and Parallel Dialogue: Knowledge Construction in Professional Conversations." *Teacher and Teaching: Theory and Practice* 12 (1): 13–31. doi:10.1080/13450600500364547.
- Ovando, M. N., and B. M. Harris. 1993. "Teachers' Perceptions of the Post-Observation Conference: Implications for Formative Evaluation." *Journal of Personnel Evaluation in Education* 7: 301–310.
- Ross, A., and C. D. Bruce. 2007. "Teacher Self-Assessment: A Mechanism for Facilitating Professional Growth." *Teaching and Teacher Education* 23: 146–159. doi:10.1016/j.tate.2006.04.03.
- Schep, M., C. Van Boxtel, and J. Noordegraaf. 2015. "Wat je kunt leren van een educatieve museumrondleiding [What You Could Learn During an Educational Guided Tour]." *Cultuur+Educatie* 42: 33–54.
- Schep, M., C. Van Boxtel, and J. Noordegraaf. 2018. "Competent Museum Guides: Defining Competencies for Use in Art and History Museums." *Museum Management and Curatorship* 33 (1): 2–24. doi:10.1080/09647775.2017.1387590.
- Sergiovani, T. J., and R. J. Starratt. 2002. *Supervision: A Redefinition*. New York: McGraw-Hill.
- Shulman, L. S. 1987. "Knowledge and Teaching: Foundations of the New Reform." *Harvard Educational Review* 57 (1): 1–23.
- Struyk, L. R., and L. H. McCoy. 1993. "Self-Evaluation Techniques." *American Secondary Education* 22 (2): 2–5.
- Sweeney, J. 1983. "The Post-Observation Conference: Key to Teacher Improvement." *The High School Journal* 66 (2): 135–140.
- Tang, S. Y. F., and A. W. K. Chow. 2007. "Communicating Feedback in Teaching Practice Supervision in a Learning-Oriented Field Experience Assessment Framework." *Teaching and Teacher Education* 23 (7): 1066–1085. doi:10.1016/j.tate.2006.07.013.
- Vasquez, C. 2004. "'Very Carefully Managed': Advice and Suggestions in Post-Observation Meetings." *Linguistics and Education* 15: 33–58. doi:10.1016/j.linged.2004.10.004.
- Veenman, S., and E. Denessen. 2001. "The Coaching of Teachers: Results of Five Training Studies." *Educational Research and Evaluation* 7 (4): 385–417. doi:10.1076/edre.7.4.385.8936.
- Waite, D. 1995. *Rethinking Instructional Supervision: Notes on Its Language and Culture*. London: The Falmer Press.
- Walsh, S. 2011. *Exploring Classroom Discourse: Language in Action*. New York, NY: Routledge.
- Walsh, S. 2013. *Classroom Discourse and Teacher Development*. Edinburgh, UK: Edinburgh University Press.
- Waring, H. Z. 2017. "Going General as a Resource for Doing Advising in Post-Observation Conferences in Teacher Training." *Journal of Pragmatics* 110: 20–33. doi:10.1016/j.pragma.2017.01.009.

Appendix A. Total number of words per participant and participant's proportion of talk per post-observation conversation.

Conversation	Participant	Total words per conversation	Proportion of talk per conversation (%)
1	Emily	3396	51
	Anna	3300	49
2	John	1836	46
	Iris	2156	54
3	Lucy	4748	61
	Sophie	3208	39
4	Rose	2950	56
	Joanna	2320	44
5	Lucia	2149	37
	Claire	3733	63
6	Martha	3407	56
	Eliza	2702	44
7	Mary	2551	43
	David	3433	57