Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 42 (2020) 109–117

Transportation Research Procedia

www.elsevier.com/locate/procedia

46th European Transport Conference 2018, ETC 2018

Objective and subjective predictors of perceived cleanliness in train stations

Martijn C. Vos[a,b,c,*], Mark van Hagen[b]

[a]*Hanze University of Applied Sciences, school of Facility Management, Groningen, the Netherlands,*
[b]*Netherlands Railways (NS), Utrecht, the Netherlands*
[c]*University of Twente, department of Communication Science, Enschede, the Netherlands*

Abstract

Cleanliness is one of the key determinants of overall customer satisfaction in train stations. Customers' perception of cleanliness is not limited to cleaning only but depends on multiple predictors. A better understanding of these predictors may contribute to the optimisation of perceived cleanliness in train stations. The current study was designed to examine how objective predictors (measures of cleaning quality), subjective predictors (e.g., customers' perception of lighting, scent, staff), and demographic variables relate to perceived cleanliness in train stations. Data on cleaning quality were gathered by trained cleaning inspectors and data on subjective predictors of cleanliness were obtained through surveys collected at 25 train stations in the Netherlands (N = 19.206). Data were examined using correlation and regression analysis. Positive and significant correlates of perceived cleanliness in train stations were found, including: perception of scent, lighting, colour, and staff. In regression analysis, customers' perception of scent and lighting appeared to be powerful predictors of perceived cleanliness. These findings underline that customers' perception of cleanliness is not only influenced by cleaning quality, but also by other predictors, such as scent, lighting, colour, and staff behaviour.

Keywords: perceived cleanliness, predictors, train stations, customer satisfaction, customer experience, scent, lighting

1. Introduction

Changing traveller's mode of travel from car and plane to rail is seen as a key strategy of the European Union to reduce CO2 emissions (e.g., Brons, Givoni, & Rietveld, 2009). This strategy has been embraced by train operating companies (TOC's), who aim to attract more customers by improving the quality of their services. From previous research, we know that customers' perception of service quality is shaped by multiple dimensions, such as the

* Corresponding author:
E-mail address: martijn.vos@ns.nl

10.1016/j.trpro.2019.12.011

reliability of the service, responsiveness of service employees, but also the cleanliness of the service environment (Parasuraman, Zeithaml, & Berry, 1988). The latter is considered to be one of the key predictors of customer satisfaction and customers' perception of service quality (van Lierop, Badami, & El-Geneidy, 2017; Wakefield & Blodgett, 1996).

In literature, cleanliness is approached from two distinct perspectives: cleaning quality and customers' perception of cleanliness. Cleaning quality is defined as the absence of dirt (i.e., dust, stains & litter) and is mostly monitored by trained inspectors through predetermined indicators, such as the cleanliness of benches, banisters, and floors (Sherlock, O'Connell, Creamer, & Humphreys, 2009). Cleaning quality represents the objective organisational perspective on cleanliness and is mostly used by TOC's (and service organisations in general) to monitor the quality of the cleaning process. As opposed to cleaning quality, customers' perception of cleanliness is considered to be more subjective and based on information that customers capture through their senses, mostly measured through questionnaires or more interactive methods. Knowledge and standards on how to manage cleaning quality are widely available in literature (Sherlock et al., 2009; Van Ryzin, Immerwahr, & Altman, 2008) and practice (e.g., the Netherlands: NEN 2075, international: ISO 9001). Research into predictors of customers' perception of cleanliness however, is mostly explorative and qualitative (Whitehead, May, & Agahi, 2007). Therefore, quantitative empirical studies are warranted to gain deeper insight into the predictors of perceived cleanliness. Improved insight in this area will further advance the scientific knowledge base and allow TOC's to make well-informed decisions with regard to perceived cleanliness.

Hence, with this study, we aim to identify predictors of perceived cleanliness in literature and subsequently verify which variables can be used to predict perceived cleanliness at (Dutch) train stations.

2. Literature review

Following previous research, predictors of perceived cleanliness can be summarised under three broad headings, including: (1) cleaning quality, (2) appearance of the environment, and (3) staff behaviour (Vos, Galetzka, Mobach, Van Hagen, & Pruyn, 2018a; Whitehead et al., 2007). Cleaning quality is defined as the absence of dirt, such as dust, stains, and litter (Van Ryzin et al., 2008; Whitehead et al., 2007). Following the work of Bitner (1992), the appearance of the environment comprises three dimensions: ambient conditions, space/function, and signs, symbols, & artefacts. Staff behaviour is about the presence, appearance, attitude, and behaviour of employees who serve the primary (e.g., train conductor) and secondary service process (e.g., cleaning staff) (Vilnai-Yavetz & Rafaeli, 2011; Whatley, Jackson, & Taylor, 2012). In the current study, we will use this typology and identify individual predictors of perceived cleanliness that are related to cleaning quality, the appearance of the environment, and staff behaviour. See Figure 1 for our conceptual model.

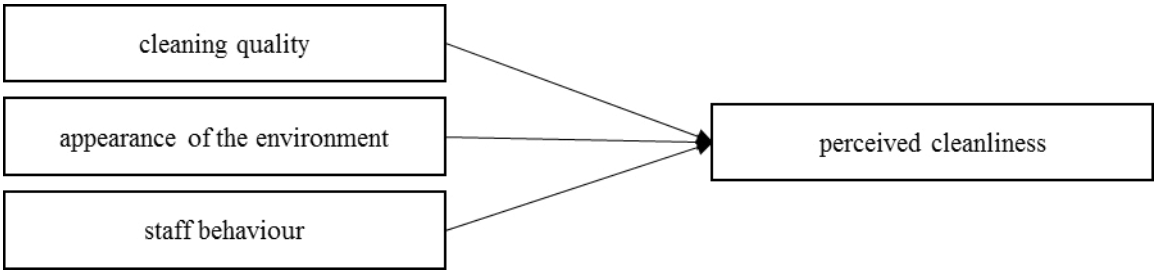

Figure 1. Conceptual framework, predictors of perceived cleanliness

2.1. Cleaning quality

The first predictor of perceived cleanliness we have identified, is cleaning quality. Literature indicates that higher levels of cleaning quality are associated with positive perceptions of cleanliness (Edgcumbe, 2009; Van Ryzin et al., 2008). Van Ryzin et al. (2008) found that objective measures of the cleanliness of the streets in New York City were correlated with citizen ratings of street cleanliness. In a hospital setting, Edgcumbe (2009) found that the patients' perceptions cleanliness were correlated with rates of Meticilline Resistente Staphylococcus aureus (MRSA)

bacteremia – a bacterium that is resistant to antibiotics. Similar findings were done in two qualitative studies (Vos et al., 2018b; Whitehead et al., 2007). Accordingly, we hypothesize (hypothesis 1) that cleaning quality is positively associated with perceived cleanliness.

2.2. Appearance of the environment

2.3. The appearance of the environment was operationalised following the environmental dimensions of the servicescapes framework (Bitner, 1992). In the following, individual predictors of perceived cleanliness, related to ambient conditions, space/function, and signs, symbols, & artefacts will be discussed.

2.4. Ambient conditions are background characteristics of the environment such as sound and temperature (e.g., Mattila & Wirtz, 2001). As a general rule of thumb, ambient conditions are perceived unconsciously and affect the five senses (Bitner, 1992). In the current study, scent, lighting, and colour were identified as ambient conditions that function as predictors of perceived cleanliness (Vos et al., 2018a, 2018b; Whitehead et al., 2007). In several studies, scent was identified as a powerful predictor of perceived cleanliness (De Lange, Debets, Ruitenburg, & Holland, 2012; Whatley et al., 2012). The absence of strong scents and/or the presence of scents that are associated with cleanliness (i.e., citrus) will lead to more positive perceptions of cleanliness. Hence, we hypothesize (hypothesis 2) that perceptions of scent are positively associated with perceived cleanliness. Lighting was identified as a predictor of perceived cleanliness in a virtual reality (VR) study (Vos & Van Hagen, 2015). More favourable perceptions of lighting and more specifically higher light intensities, were associated with more positive perceptions of cleanliness. Hence, we hypothesize (hypothesis 3) that perceptions of lighting are positively associated with perceived cleanliness. Colour was identified as the final ambient condition that functions as a predictor of perceived cleanliness. In several studies, lighter colours (e.g., white, light blue, & light green) were associated with concepts such as purity, nature, and cleanliness (Brotherton, 2007; Harutyunyan, 2015). Therefore, we predict (hypothesis 4) that perceptions of colour are positively associated with perceived cleanliness.

2.5. Space layout and functionality refers to the ways in which equipment and furnishing are arranged, the size and shape of those items, and the spatial relationship among them. Previous literature has identified crowding as a predictor of perceived cleanliness (Vos et al., 2018b; Whitehead et al., 2007). Crowding refers to the psychological feeling of not possessing available space (Vos et al., 2018b) and is very often dependent on the spatial layout and functionality of an environment (Stokols, 1972). Evidence is scarce, and it remains unclear how crowding affects perceived cleanliness. We found two possible explanations. First, density is positive as it covers traces of uncleanliness, but hinders the efficiency of the cleaning process. Second, density is negative as it relates to crowdedness (e.g., irritation) and the idea that other people are a (potential) source of litter, diseases, and unpleasant odours. The latter explanation was considered to be most plausible, hence, we expect (hypothesis 5) that density is negatively associated with perceived cleanliness.

2.6. Signs, symbols, & artefacts function as explicit or implicit signals that communicate about rules or the image of the service organisation to its users. In the current study, no predictors of perceived cleanliness related to signs, symbols, & artefacts were identified.

2.7. Staff behaviour

The interaction between customers and service personnel (i.e., the service encounter) is considered to be a crucial factor in the service delivery process which influences perceived cleanliness (Vos et al., 2019). Vos et al. (2019) found that train passengers had positive perceptions of cleanliness when cleaning staff was present – compared to a condition without cleaning staff. Whitehead et al. (2007) found similar effects for staff that serves the primary process and is not primarily involved in the cleaning process. Therefore, we hypothesize (hypothesis 6) that perceptions of staff are positively associated with perceived cleanliness.

2.8. Demographics

Two basic demographic variables were included in this study: gender and age. Literature on gender and perceived cleanliness shows inconsistent findings. Barber & Scarcelli (2010) and Vos et al. (2019) found no differences

between male and female participants with respect to perceived cleanliness. Similarly, Zemke, Neal, Shoemaker & Kirsch (2015) and Lockyer (2003) reported no significant differences between male and female on perceived cleanliness, although female participants were more concerned about cleanliness in both studies. Mortimer & Clarke (2011) and dell'Olio, Ibeas & Cecin (2011) also found that cleanliness was more important for female than for male. Since previous research on gender differences is inconclusive, we expect (hypothesis 7) that gender is not associated with perceived cleanliness.

Literature on age as a predictor of perceived cleanliness appeared to be more conclusive. Zemke et al. (2015) and Barber & Scarcelli (2010) found no differences between age groups in ratings of perceived cleanliness. Hence, we expect (hypothesis 8) that age is not associated with perceived cleanliness.

3. Method

3.1. Sample

The current study was cross-sectional; questionnaires were administered to passengers once (N = 19.206). Data were collected on 25 largest Dutch train stations (based on number of passengers per day) during the year 2017. The included stations are located in different Dutch cities, either in urban or suburban environment. All stations were operated by the same TOC and were fairly similar with respect to the number of passenger. In addition, cleaning frequencies on all train stations were equal as well. Apart from these similarities, there were some situational differences between the 25 train stations (e.g., geographical location, architectural design) that were beyond the scope of this study and for which we controlled in our statistical analysis. These stations are located in different Dutch cities, either in an urban or suburban environment. Descriptive information about the sample is presented in Appendix 1.

3.2. Measures

Cleaning quality was monitored through visual inspection by experienced inspectors using predetermined indicators. Predetermined indicators of cleaning quality for example included the number of stains on banisters or flooring. Cleaning quality data were collected monthly, as part of the regular cleaning process with the main aim of evaluating the performance of the cleaning department. The measure represents the average value of the scores given by the inspectors to the predetermined indicators of cleaning quality with response values ranging from one to ten.

Customers' perceptions were measured using an adapted version of the 'station experience monitor' (Van Hagen, 2015). Customers' perceptions of cleanliness, scent, lighting, colour, overview, staff, and density were assessed using a single item with response value ranging from one to ten, with higher values referring to more positive perceptions. See Appendix 2 for the complete questionnaire. Measures of cleaning quality and customers' perceptions were collected in the same week.

3.3. Statistical analysis

Descriptive statistics of perceived cleanliness were explored by inspection of the mean and standard deviation. A correlation matrix was constructed in order to identify correlates of perceived cleanliness and verify our hypotheses. Potential differences in perceived cleanliness score for gender were examined using a two-sided t-test. To identify useful predictors of perceived cleanliness at train stations, we used regression analysis. A regression model was developed using a stepwise forward model selection procedure – a procedure frequently used in studies with an explorative character (Henderson & Denison, 1989). At each step a new predictor was added to the minimal model which explained the most variance in the outcome variable that was unexplained by the variables already included in the model.

Examination of the residuals of the final regression model showed no violations of independence, normality, homogeneity, or linearity assumptions. No violation of regression assumptions was found, indicating that model errors are uncorrelated between passengers even though passengers were sampled nested within train stations.

4. Results

4.1. Descriptive statistics

Table 1 presents for each continuous variable the mean, standard deviation, and Pearson correlation coefficient with the other variables (** $p < .01$). Perceived cleanliness was rated 7.08 on average, with individual ratings ranging from 1 (n = 163) till 10 (n = 1,243). Further examination of the results showed that 83.5% evaluated cleanliness with a 6 or higher. Meaning that passengers are on average reasonably satisfied according to the grading system that is commonly used in schools in the Netherlands, in which 5.5 marks the distinction between "pass" and "fail".

Table 1. Descriptive statistics for continuous variables

	M	*SD*	Correlation coefficients							
Variable			**1.**	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**	**8.**
1. Perception cleanliness	7.08	1.72	-							
2. Cleaning quality	7.00	1.95	.06**	-						
3. Perception scent	6.83	1.75	.75**	.05**	-					
4. Perception lighting	7.47	1.58	.57**	.03**	.51**	-				
5. Perception colour	5.65	1.93	.39**	.06**	.41**	.36**	-			
6. Perception staff	7.57	1.57	.43**	.00	.42**	.49**	.27**	-		
7. Perception density	6.48	1.78	.00	-.02**	-.05	.06**	.08**	.05**	-	
8. Age	32.2	16.4	.03**	.01	.00	-.02	.11**	.06**	.06**	-

** $p < .01$

The correlation coefficients (Table 1) indicate how objective and subjective predictors are related to perceived cleanliness and each other. Passengers' perception of scent turned out to be strongest correlate of perceived cleanliness of train stations.

Our sample contains 56% female and 44% male passengers. The results of a two-sided t-test suggest that perceived cleanliness is not perceived different by female (M = 7.06, SD = 1.71) and male (M = 7.10, SD = 1.73) passengers (t = 1.81, df = 18,455, p = .07).

4.2. Returning to the hypotheses

The correlation coefficients (see Table 1, column '1') indicate that empirical support was found for five of our hypotheses: perceived cleanliness is positively and significantly associated with the perception of scent (hypothesis 2), perception of lighting (hypothesis 3), perception of colour (hypothesis 4), staff (hypothesis 6) and not associated with gender (hypothesis 7) and age (hypothesis 8). Following the correlation coefficients, the effect sizes of the associations can be regarded as large for the perception of scent and lighting, medium to large for the perception of staff and colour, and small for age (Cohen, 1992). No empirical evidence was found for our hypothesis regarding cleaning quality (hypothesis 1) and perceived density (hypothesis 5).

4.3. Regression analysis

To identify useful predictors of perceived cleanliness at train stations, a regression model was developed using a stepwise forward model selection procedure. The station on which data was collected, was included as a control variable in the minimal model after an ANOVA F-test showed that perceived cleanliness differed significantly between the train stations in our sample (F = 58,33, df = 24, $p < .001$). At each step a new predictor was added to the minimal model which explained the most variance in the outcome variable that was unexplained by the variables already included in the model. The order in which the predictors were added, was shown in Table 2. This table indicates the amount of remaining variance the newly added variable explains on top of what is already explained by the variables already included in the model. Through the stepwise selection procedure, we have identified a

regression model that includes the train stations as control variable, customers' perceptions of scent, lighting, staff, and colour. We have summarized the results of this model in Table 3. The table presents the estimates, standard errors, confidence intervals, and p-values for each predictor, station 1 is the baseline with which the other stations are compared.

Table 2. Steps model selection

Predictors	**Adjusted R^2 of the model after each step**
Minimal model: train stations	.06
Step 1: perception scent	.58
Step 2: perception lighting	.63
Step 3: perception staff	.63
Step 4: perception colour	.63

Table 3. Results final regression model

Predictors	**Estimate (β)**	**Standard Error**	**95% confidence interval**	***p*-value**
β_0 (intercept)	.64	.07	.50 - .78	> .001
$\beta_{\text{Station 2}}$	-.28	.06	-.41 - -.16	> .001
$\beta_{\text{Station 3}}$	.16	.06	.04 - .29	.010
$\beta_{\text{Station 4}}$	.15	.06	.03 - .28	.014
$\beta_{\text{Station 5}}$	.04	.06	-.08 - .16	.492
$\beta_{\text{Station 6}}$	-.13	.06	-.25 - -.00	.041
$\beta_{\text{Station 7}}$	.02	.06	-.09 - .14	.701
$\beta_{\text{Station 8}}$	-.24	.06	-.36 - -.12	> .001
$\beta_{\text{Station 9}}$	-.29	.06	-.42 - -.16	> .001
$\beta_{\text{Station 10}}$	-.15	.06	-.27 - -.03	.014
$\beta_{\text{Station 11}}$	.14	.06	.02 - .26	.018
$\beta_{\text{Station 12}}$	-.07	.06	-.19 - .04	.231
$\beta_{\text{Station 13}}$	-.05	.06	-.19 - .07	.385
$\beta_{\text{Station 14}}$	.06	.06	-.06 - .18	.345
$\beta_{\text{Station 15}}$	-.00	.06	-.13 - .11	.893
$\beta_{\text{Station 16}}$	-.17	.06	-.29 - -.05	.006
$\beta_{\text{Station 17}}$	-.01	.06	-.14 - .10	.761
$\beta_{\text{Station 18}}$	-.08	.06	-.20 - .03	.173
$\beta_{\text{Station 19}}$	-.10	.06	-.22 - .02	.105
$\beta_{\text{Station 20}}$	.42	.06	.29 - .55	> .001
$\beta_{\text{Station 21}}$	.25	.06	.13 - .37	> .001
$\beta_{\text{Station 22}}$	-.06	.06	-.18 - .06	.331
$\beta_{\text{Station 23}}$	-.15	.06	-.28 - -.02	.016
$\beta_{\text{Station 24}}$	-.16	.06	-.29 - -.03	.010
$\beta_{\text{Station 25}}$	-.09	.06	-.21 - .03	.153
$\beta_{\text{Perception scent}}$	.56	.00	.55 - .58	> .001
$\beta_{\text{Perception lighting}}$	.23	.00	.21 - .24	> .001
$\beta_{\text{Perception staff}}$	.06	.00	.05 - .07	> .001
$\beta_{\text{Perception colour}}$	.06	.00	.05 - .07	> .001

The final model explains 63.7% of the variance in perceived cleanliness in train stations (R^2_{Adj} = .637). The largest amount of variance was explained by customers' perception of scent, the train station and customers' perception of lighting. Customers' perception of staff and colour added marginal predictive power to the model.

Based on the regression results, we conclude that, for a given train station, customers' perceptions of scent and lighting are important predictors of their perception of cleanliness.

5. Discussion

5.1. Interpretation and explanation of the results

The current study identified customers' perceptions of scent, lighting, staff, and colour as significant (positive) correlates of customers' perception of cleanliness in train stations. No empirical support was found for the hypothesized association with cleaning quality, customers' perception of density, and the demographic variables gender and age. Customers' perception of scent and lighting appeared to be the most powerful predictors of differences in perceived cleanliness in train stations.

The association with customers' perception of scent, lighting, staff, and colour clearly confirms pervious findings indicating that perceived cleanliness is more comprehensive than cleaning quality only (Vos et al., 2018b; Whitehead et al., 2007). Apparently, the presence of pleasant environmental cues, such as scent, imply cleanliness of the environment. This phenomenon can be interpreted following the 'halo effect' (Eagly, Ashmore, Makhijani, & Longo, 1991) which suggests that customers tend to assume that one positive trait (i.e., pleasant scents, helpful staff), imply the presence of other positive traits (e.g., clean environment).

Measures of cleaning quality did show a significant but very weak correlation (.06) with perceived cleanliness (see: Table 1). Which may imply that the predetermined indicators of cleaning quality, observed by trained inspectors, such as cleaning quality of benches or flooring, are not the indicators that customers use to determine if a train station is clean. In general, customers may have little interest in cleaning quality and will pay little attention to cleanliness if quality is above a minimum level – which applies to most (Dutch) train stations. As a result, we expect that cleaning quality mainly affects perceived cleanliness when excessively bad or remarkably good (Stipak, 1979).

Customers' perception of density does not seem to be associated with perceived cleanliness. Apparently, passengers do not associate density with perceived cleanliness. In line with our expectations with regard to age and gender, neither of the demographic variables were associated with perceived cleanliness.

5.2. Limitations and perspectives for further research

The size of our data set, and its combination of data from different train stations, enabled us to show hypothesis-driven associations between perceived cleanliness and its predictors. As a wide range of train stations in the Netherlands were included, results are considered to be representative for the Netherlands and similar countries. It would be interesting to verify our findings in future research, by including data from other public transport organisations in different countries. Another strength of the current study included that data on objective and subjective measures of cleanliness were collected at the same point in time.

Survey-based studies, like the current one, are limited with regard to the number of predictors that can be included. The search for predictors may be broadened by including new predictors, such as architectural characteristics of train stations. But the search may also be deepened, by specifying which specific scent, light setting, staff behaviour, or colour may enhance customers' perception of cleanliness. A field experiment approach seems most appropriate to produce such insights.

5.3. Implications for practice

Probably the most important implication for practice is that TOC's should reconsider the way in which perceived cleanliness is currently managed. Besides keeping cleaning quality on a certain level, the appearance of the environment and staff behaviour should be managed accordingly. Station managers could for example think about adding pleasant scents to the platforms or bright lighting and bright but cool colours (e.g., blue, green) to tunnels (van Hagen, 2011), and increase the visibility of the cleaning process by introducing day-time cleaning. Such actions might have a substantial impact on customers' perception of cleanliness and subsequently their satisfaction and perception of overall service quality.

References

Barber, N., & Scarcelli, J. M. (2010). Enhancing the assessment of tangible service quality through the creation of a cleanliness measurement scale. *Managing Service Quality, 20*(1), 70-88.

Bitner, M. J. (1992). Servicescapes: The impact of physical surroundings on customers and employees. *Journal of Marketing, 56*(2), 57-71.

Brons, M., Givoni, M., & Rietveld, P. (2009). Access to railway stations and its potential in increasing rail use. *Transportation Research Part A: Policy and Practice, 43*(2), 136-149.

Brotherton, B. & Wood, R. C. (2007). *The Sage Handbook of Hospitality Management.* London, UK: Sage Publications.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155.

Cordera, R., Coppola, P., Dell'Olio, L., Ibeas, A. (2017) 'Is accessibility relevant in trip generation? Modelling the interaction between trip generation and accessibility taking into account spatial effects'. Transportation, Vol. 44 (6), pp 1577–1603

Dell'Olio, L., Ibeas, A., & Cecin, P. (2011). The quality of service desired by public transport users. *Transport Policy, 18*(1), 217-227.

De Lange, M. A., Debets, L. W., Ruitenburg, K., & Holland, R. W. (2012). Making less of a mess: Scent exposure as a tool for behavioral change. *Social Influence, 7*(2), 90-97.

Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but…: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin, 110*(1), 109.

Edgcumbe, D. P. (2009). Patients' perceptions of hospital cleanliness are correlated with rates of meticillin-resistant Staphylococcus aureus bacteraemia. *Journal of Hospital Infection, 71*(1), 99-101.

Harutyunyan, K. (2015). Colour terms in advertisement. *Armenian Folia Anglistika*, *2*(14), 56-67.

Henderson, D. A., & Denison, D. R. (1989). Stepwise regression in social and psychological research. *Psychological Reports, 64*(1), 251-257.

Lockyer, T. (2003). Hotel cleanliness—how do guests view it? Let us get specific. A New Zealand study. *International Journal of Hospitality Management, 22*(3), 297-305.

Mattila, A. S., & Wirtz, J. (2001). Congruency of scent and music as a driver of in-store evaluations and behavior. *Journal of Retailing, 77*(2), 273-289.

Mortimer, G., & Clarke, P. (2011). Supermarket consumers and gender differences relating to their perceived importance levels of store characteristics. *Journal of Retailing and Consumer Services, 18*(6), 575-585.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). Servqual: A multiple-item scale for measuring consumer perc. *Journal of Retailing, 64*(1), 12.

Sherlock, O., O'Connell, N., Creamer, E., & Humphreys, H. (2009). Is it really clean? An evaluation of the efficacy of four methods for determining hospital cleanliness. *Journal of Hospital Infection, 72*(2), 140-146.

Stipak, B. (1979). Citizen satisfaction with urban services: potential misuse as a performance indicator. *Public Administration Review, 39*(1), 46-52.

Stokols, D. (1972). On the distinction between density and crowding: some implications for future research. *Psychological Review, 79*(3), 275.

van Hagen, M. (2011). *Waiting experience at train stations*. Delft, Nederland: Eburon Uitgeverij BV.

van Hagen, M. (2015). Effect of station improvement measures on customer satisfaction. *Journal of Traffic and Transportation Engineering*, *3*(1), 7-18.

Van Lierop, D., Badami, M. G., & El-Geneidy, A. M. (2017). What influences satisfaction and loyalty in public transport? A review of the literature. *Transport Reviews*, 1-21.

Van Ryzin, G. G., Immerwahr, S., & Altman, S. (2008). Measuring street cleanliness: A comparison of New York City's scorecard and results from a citizen survey. *Public Administration Review, 68*(2), 295-303.

Vilnai-Yavetz, I., & Rafaeli, A. (2011). The effects of a service provider's messy appearance on customer reactions. *Services Marketing Quarterly, 32*(3), 161-180.

Vos, M., & Van Hagen, M. (2015). Licht op treinstations (light on train stations). *In Presentation at Colloquium Vervoersplanologisch Speurwerk (CVS), November 19-20, 2015, Antwerp, Belgium.*

Vos, M. C., Galetzka, M., Mobach, M. P., Van Hagen, M., & Pruyn, A. T. H. (2018a). Cleanliness unravelled: A review and integration of literature. *Journal of Facilities Management, 16*(4).

Vos, M. C., Galetzka, M., Mobach, M. P., Van Hagen, M., & Pruyn, A. T. H. (2018b). Exploring cleanliness in the Dutch facilities management industry: a Delphi approach. *Facilities, 36*(9/10).

Vos, M. C., Sauren, J., Knoop, O., Galetzka, M., Mobach, M. P., & Pruyn, A. T. (2019). Into the light: effects of the presence of cleaning staff on customer experience. *Facilities*, *37*(1/2), 91-102.

Wakefield, K. L., & Blodgett, J. G. (1996). The effect of the servicescape on customers' behavioral intentions in leisure service settings. *Journal of Services Marketing, 10*(6), 45-61.

Whatley, V., Jackson, L., & Taylor, J. (2012). Improving public perceptions around cleanliness and health care associated infection in hospitals (service improvement). *Journal of Infection Prevention, 13*(6), 192-199.

Whitehead, H., May, D., & Agahi, H. (2007). An exploratory study into the factors that influence patients' perceptions of cleanliness in an acute NH'S trust hospital. *Journal of Facilities Management, 5*(4), 275-289.

Zemke, D. M. V., Neal, J., Shoemaker, S., & Kirsch, K. (2015). Hotel cleanliness: will guests pay for enhanced disinfection? *International Journal of Contemporary Hospitality Management, 27*(4), 690-710.

Appendix 1: Descriptive information included train stations

Train station	Train station size*	N	Age				% female
			Mean	SD	Min.	Max.	
1.	Very large	820	33.39	16.02	11	84	56%
2.	Very large	749	33.71	16.23	12	81	52%
3.	Very large	716	33.90	16.20	12	82	55%
4.	Very large	666	37.29	16.61	13	81	56%
5.	Large	710	37.06	16.69	14	82	54%
6.	Large	777	30.43	15.79	12	83	54%
7.	Large	883	30.85	17.29	12	86	55%
8.	Large	780	29.08	15.70	14	83	56%
9.	Large	745	35.09	16.74	12	85	57%
10.	Large	781	29.38	15.09	14	81	62%
11.	Large	808	31.99	16.74	12	84	56%
12.	Large	771	29.67	16.61	12	81	52%
13.	Large	751	35.15	15.06	13	85	52%
14.	Large	750	32.20	17.38	11	85	53%
15.	Large	808	32.30	17.08	12	86	53%
16.	Large	778	35.43	17.64	12	87	55%
17.	Large	777	29.55	13.37	12	75	53%
18.	Large	764	31.42	17.75	13	90	57%
19.	Large	816	28.71	14.47	13	87	55%
20.	Large	743	32.52	16.53	13	89	49%
21.	Large	738	30.79	17.22	12	82	55%
22.	Medium	782	29.77	16.51	12	84	58%
23.	Medium	789	33.75	15.51	14	85	56%
24.	Medium	711	31.65	16.22	12	83	55%
25.	Medium	793	32.47	16.26	12	83	56%
Total sample		19,206	32.30	16.27	11	90	56%

** Size train station: 80,000 or more passengers per day = very large; 79,999 – 30,000 = large; 29,999 or less = medium*

Appendix 2: Overview of questionnaire items and answer types

Question	Answer type
1. Gender	male / female
2. Age	years
3. I think this station is clean	ten-point scale*
4. I think the scent on this station is pleasant	ten-point scale*
5. I think the lighting on this station is pleasant	ten-point scale*
6. I think this station is colourful	ten-point scale*
7. I think the staff on this station is friendly	ten-point scale*
8. I think this station is crowded	ten-point scale*

** ten-point scale: (1) totally disagree (2) (3) (4) (5) (6) (7) (8) (9) (10) totally agree*