



Repositorio Institucional de la Universidad Autónoma de Madrid

<https://repositorio.uam.es>

Esta es la **versión de autor** del artículo publicado en:

This is an **author produced version** of a paper published in:

Spanish Journal of Psychology 19 (2016): e39

DOI: <http://dx.doi.org/10.1017/sjp.2016.39>

Copyright: © Universidad Complutense de Madrid and Colegio Oficial de Psicólogos de Madrid

El acceso a la versión del editor puede requerir la suscripción del recurso
Access to the published version may require subscription

Re-paying Attention to Visitor Behavior: a re-analysis using meta-analytic techniques

Re-paying Attention

Fecha de envío: 29-09-2015

Resumen

Aunque la actual acumulación de Estudios de público en espacios patrimoniales revela el reconocimiento de su importancia, carecemos de estudios acumulativos de resultados. Este trabajo, basado en el empleo de las técnicas meta-analíticas al campo de investigación de los museos, reanaliza con dichas técnicas la base de datos de Serrell (1988), dando como fruto una serie de variables de medición en el ámbito del uso del espacio expositivo, que van más allá de las variables clásicas. Para esta base de datos, que incluye 110 estudios, se han definido cuatro índices de tamaño del efecto y se han analizado el papel de variables moderadoras relevantes. Las principales conclusiones indican que el tiempo de visita por cada elemento se relaciona con el tamaño de la exposición y con su carácter novedoso, y que la velocidad de desplazamiento es mayor cuanto mayor es el espacio que ocupa. Los índices obtenidos en este trabajo, pueden entenderse como referentes de comparación con los resultados obtenidos en nuevas evaluaciones y pueden ayudarnos a predecir el funcionamiento de una nueva exposición, dimensionando los problemas de forma tangible y aportando al campo nuevas herramientas de evaluación más actuales y potentes.

Palabras clave: estudios de público, evaluación en museos, estudios en museos, meta-análisis.

Abstract

Although current visitor studies accumulation reveals the recognition of its importance, actually there are not cumulative outcome studies. This work, is based on the use of meta-analytic techniques in the research field of museums. With such techniques we analyzed the database Serrell (1988), obtaining as a result a number of variables

measuring the level of use exhibition space, which go beyond the classical variables. For this database, which includes 110 studies, have defined four levels of effect size and analyzed the role of relevant moderating variables. Key findings indicate that the visiting time for each element relates to the size of the exhibition and its newness, and the speed is greater the greater the space it occupies. The indices obtained in this work can be understood as a reference for comparison with the results of new evaluations and may help predict the performance of a new exhibition, sizing problems tangibly and providing new tools more powerful current assessment to research field.

Keywords: visitor studies, museum studies, meta-analysis.

Visitor Studies examine the fit between different visitor-group segments and museum offerings (Klingler & Graft, 2012). These studies were consolidated in the 1980s with the work of psychologists such as Chan Screven, Harris Shettell, and Ross Loomis (Bitgood, 2011). These studies make part of a generalized design focus centered on users (Abrams, Maloney-Krichmar & Preece, 2004) and focus on studying visitor behavior at exhibits using multiple levels of analysis (Asensio & Asenjo, 2011; Heath & Lehn, 2010). Understanding visitor processes and their interactions with exhibits requires to evaluate the exhibit on display and thus be able to predict and design future exhibits and new activities (Falk & Dierking, 2013).

Classical museums are based on exhibits made of elements that give precedence to a disciplinary discourse, normally encrypted at a high conceptual level with few communicative aims (Carbonell, 2012; Knell et al., 2011). New exhibits, supposedly based on more-powerful communicative mediators, have raised the bar concerning concepts for the users to comprehend (MacDonald, 2006; Weil, 2002). Visitor Studies are increasingly more necessary (Daignault, 2011); evaluation and planning studies aspire to a minimum of 5% of the total project budget (Harlow, 2014).

The Serrell Study

Despite the great number of visitor studies, few exist that are accumulative or comparative. The museum evaluation field is a restricted area for publishing. It is not possible to find primary studies with the original data because the comprehensive evaluations have both positive and negative results. Those affect the institution image and they include confidentiality compromises. However, in 1998 Serrell published a study, *Paying attention: Visitors and Museum Exhibition*, where they got databases from more of one hundred studies directly from the evaluators in those studies. It offers

a unique source to find quantitative references for comparing the results of new studies and for establishing quantitative goals of interventions in the field.

Our study takes as its starting point the report by Serrell. That study assesses five variables to compare the results for time spent and stops made in exhibition spaces in a broad set of United States museums. The authors generated two indexes for comparing data between studies. The first, *Sweep Rate Index (SRI)*, is the total median time used by visitors, divided by the square footage of the exhibit. Because it divides time by space, it involves the *inverse of velocity (Iv)*. The second, *Percentage of Diligent Visitors (%DV)*, is the percentage of visitors who stop by at least half of the exhibit features. The authors analyzed the data with descriptive statistics, standard models of ANOVA and correlations.

Advantages of reanalyzing data with meta-analytic techniques

The term meta-analysis was coined Glass (1976) to refer to the methodology conceived for statistically analyzing and synthesizing an original set of statistical results from a sample of studies with a certain degree of homogeneity. The most significant components in developing the meta-analytical methodology were observed in the pioneering work of Glass in 1976 (Botella & Sánchez-Meca, 2015), but main advances in statistical methods came from the contributions of Hedges and Olkin (1985).

Meta-analyses consider effect size (ES) values. When the ES is calculated for each study, the measures are transformed to standardized measures allowing direct comparison between studies. The ES quantifies the relevance of an effect obtained within a particular field of study, helping us clarify if a statistically significant effect is actually relevant. For the variables of interest, a combined estimation of ES is

calculated, which, if applicable, allows assessing the effect of potentially moderating variables.

Goals of the present study

Our general purpose is performing a re-analysis of the database collected by Serrell employing meta-analytic techniques. This involves some specific goals. First, proposing ES indices suitable for the kind of studies synthesized, which involves adapting formulas for the variances of those indices. Second, obtaining weighted combinations of the values in the database. Third, analyzing any variability observed in the values, offering explanations by fitting models that include moderator variables.

Having estimates of these indices will be a significant progress in visitors' studies. On the one hand, the results of new studies will be compared with reference values; secondly, when designing new exhibitions the actual behavior of real visitors estimated from a large sample of studies can be taken in account. Furthermore, the results of the analysis of the moderating variables allow that both the reference values and the estimates used in the design can be adjusted for the characteristics of each new exhibition.

Methods

Studies included

The source of studies is the compilation published in the Serrell (1998) report. It includes data on exposition space usage in 110 exhibits in 62 different museums. The authors include studies performed on exhibits with a concrete expository message following a criterion of suitability and accessibility, for a significant sampling of museums. Each study has a sample of 40 or more randomly selected visitors. The data-

gathering method was non-participant observation, recording data related with time spent and stops throughout the exhibit space (Asensio & Pol, 2005).

Eight of the studies included in the original report were excluded of our analyses (those with numbers 19, 25, 32, 47, 56, 60, 92, and 99) because there was not enough information for calculating the statistics and the indices employed. As described below, Serrell and Adam's report provided histograms of two observed variables of each primary study. They also provided the mean value of those variables on each study (in the eight studies excluded the histogram was missing). Moreover, study 41 was also excluded because there was a major incongruence between the mean time spent by visitors as calculated from the histogram and the value reported in the text of the report (see below).

Variables in the report

The Serrell report included for the primary studies the following information: the size of the sample of visitors in the study, six variables related with characteristics of the exhibition (four coded and two quantitative variables), and data of two outcome variables. The four coded variables and their distributions appear in table 1; that table also includes descriptive statistics of the sample size and the two quantitative variables. The two observed variables in the study were *time spent* and *number of stops*. For (almost) each of those variables the report included histograms built with the data of the visitors and the means.

——— Insert Table 1 about here ———

We directly assumed the reported values of the six moderators. However, as we needed the variances of the outcome variables in each study the complete frequency distributions of both variables were reconstructed from the histograms (with a millimeter rule). The distributions allowed calculate both the average and the variance of each study in both variables. As the report provided arithmetic means the coincidence between our calculated means and those in the report allowed validate this process. Thus, our final database included the mean and the variance of both observed variables in each primary study of the final set.

Effect Size Indexes

Meta-analytic techniques are applied to values of some effect size (ES) index that reflects information related to some question of interest (Kelley & Preacher, 2012). While in the majority of meta-analyses the most well-known indexes are suitable (Borenstein, Hedges, Higgins, & Rothstein, 2009; Ellis, 2010), sometimes the meta-analyst must define specific indexes that reflect better the requirements of a particular context—as is the case here. Following, we define the four ES indexes that will be analyzed, specifying also the procedure for calculating their variances, as an estimate of the variance is necessary for weighting by the inverse of the variances.

(a) *Average Time per Feature (ATF)*. It is defined as the average time spent to the items in the exhibit during the visit, in minutes. The total time in the visit is taken from the histograms. Still, these values reflect the trivial fact that visitors spend more time in larger exhibits. For this reason, we prefer to use *ATF*. If \bar{T}_j is the average time of the sample in exhibit j , which is made up of F_j features, its *ATF* is calculated as follows:

$$ATF_j = \frac{\bar{T}_j}{F_j} \quad (1)$$

The ATF_j variance is calculated by the following formula (Botella, Suero, & Ximénez, 2012), in which S_j^2 is the variance of the visit time for the sample and N_j is the number of visitors in the sample.

$$\hat{\sigma}^2(ATF_j) = \frac{S_j^2}{F_j^2 \cdot N_j} \quad (2)$$

(b) *Percentage of Diligent Visitors (dv)*. It is the percentage of sample visitors who stopped at least in 50% of the museum features. In the original review, the variable used was %DV. In order to analyze those data we first transformed the %DV values dividing them by 100 and calculating their logit to mitigate distribution problems (Newcombe, 2012). So, the values analyzed were calculated as:

$$\text{logit}(dv_j) = \ln\left(\frac{dv_j}{1-dv_j}\right) \quad (3)$$

The variance of $\text{logit}(dv)$ is given by (Newcombe, 2012):

$$\hat{\sigma}^2[\text{logit}(dv_j)] = \frac{1}{N_j \cdot dv_j \cdot (1-dv_j)} \quad (4)$$

(c) *Inverse of Velocity (Iv)*. It is the mean of the time spent by the visitors per hundred square meters (m^2) of the exhibit space. As the exhibit area was given in square feet in the original report (ft^2) the values were converted to m^2 before the analysis ($\text{m}^2 = \text{ft}^2 \cdot 0.0929$). Then, we calculated the SU value, which represents the surface in units of 100 m^2 .

The ES and Iv indexes and their variance were calculated as (\bar{T}_j and N_j were defined above):

$$Iv_j = \frac{\bar{T}_j}{SU_j} \quad (5)$$

$$\hat{\sigma}^2(Iv_j) = \frac{S_{T_j}^2}{SU_j^2 \cdot N_j} \quad (6)$$

(d) *Stops per Feature (PF)*. It is the average number of stops made by the visitors per feature in the exhibit. First, \bar{P}_j was calculated, which is the mean number of stops made by the sample (via histograms). However, for the same reason as mentioned for *ATF*, we were more interested in the average number of stops per feature, that is:

$$\overline{PF}_j = \frac{\bar{P}_j}{F_j} \quad (7)$$

The variance for this index is as follows:

$$\hat{\sigma}^2(\overline{PF}_j) = \frac{S_p^2}{F_j^2 \cdot N_j} \quad (8)$$

Where S_p^2 is the sample variance of the number of stops in study j .

Meta-analysis Procedures

Statistical analyses were performed with the procedures developed by Hedges and Olkin (1985), weighting the studies by the inverse of the variance. We assumed a random effects (RE) model for all of the analyses because it is more conservative than a fixed effect (FE) model, and allows generalizing the results beyond the specific set of

studies analyzed (Borenstein, Hedges, Higgins, & Rothstein, 2010; Hedges & Vevea, 1998; Raundenbusch, 2009). The inter-study variance was estimated by the maximum likelihood method. Assuming an RE model means that the set of studies estimate a distribution of ES in the population with $\theta_i \sim N(\mu_\theta; \tau^2)$. Each individual ES estimates a population ES, assuming that the error variability for that study is due to sampling (σ_j^2). Then, the estimated between studies variance (τ^2) is added to the variance of each study. The RE model involves applying an inverse weight of variance such that:

$$w_j = \frac{1}{(\sigma_j^2 + \tau^2)} \quad (9)$$

Because the variances are unknown, corresponding estimators are used. We have already defined the sampling variances of our four ES indexes in formulas [2], [4], [6], and [8].

Calculations were done with *METAFOR* (Viechtbauer, 2010; asymmetry tests and funnel plots). We also performed sensitivity analyses, paying attention to outliers (Higgins & Green, 2008). Those analyses confirmed that the results hardly change when the outliers were excluded. A special treatment has not been performed on missing values, given that these values belong to the studies we excluded from the meta-analysis, as explained in the following section.

To assess the heterogeneity of the estimates we used the Q and I² statistics. Q involves a significance test of whether the amount of variability observed exceeds the expected under a fixed effect model. The I² statistic allows assessing the degree of variability beyond that dichotomous decision. There are good reasons to use both indexes in combination (Huedo-Medina, Sánchez-Meca, Marín-Martínez, & Botella, 2006).

Moderating Variables

The six characteristics of the studies provided by the original report were treated as potentially moderating variables, analyzing their associations with the values in the four ES indices. As anticipated above, we directly assumed the codes offered in the Serrell report. The only exception was the conversion of square feet (ft²) to the more international square meters (m²) units, as explained above.

Results

Average time per feature (ATF)

The combined weighted estimate of the population value (ATF_{\bullet}) is 0.43 minutes per feature (CI: 0.49; 0.37). The values show a high degree of heterogeneity [$Q_w(100) = 4826.69, p < .001; I^2 = 99.53\%$ (CI: 100.13; 98.93)]. The estimate of the specific variance (τ^2) equals 0.0866 ($SE = 0.0125$). Table 2 summarizes the results obtained when fitting meta-regression models with the two quantitative moderators, and the results when fitting models with the four categorical moderators.

——— Insert Table 2 about here ———

The moderator *number of features* explains a significant percentage of the variance in the ES values, even although it is low in absolute terms. The negative sign of the slope means that the greater the number of features, the lower the *ATF* spent by visitors. The *surface* moderator is not significant.

Only one of the categorical moderators, *exhibition age*, was significant. The combined values for the categories created with that model are $ATF_{\bullet(\text{new})} = 0.45$ and $ATF_{\bullet(\text{old})} = 0.19$. Thus, in new exhibits visitors devote an average of 0.26 minutes

(15.06 seconds) more per feature than in old exhibits. We have fitted a combined model with the two moderating variables that are significant (number of features and exhibition age). It explains a significant portion (16.39%) of the observed heterogeneity [$Q_M(2) = 17.9456$; $p < .001$]; the regression equation is:

$$ATF'_i = 0.0036 - 0.0037 \cdot \text{Number of Features}_i + 0.2903 \cdot \text{Age}_i$$

According to this model, for each additional feature in the exhibit, the mean time spent by a typical visitor is reduced by 0.0037 minutes (about 0.2 seconds). By contrast, the mean time used per visitor increases 0.2903 minutes (17.41 seconds) in the case of a new exhibit, as compared to an old exhibit.

Percentage of Diligent Visitors (dv)

The combined weighted estimate (dv_\bullet), after the inverse logit transformation is completed, is 0.30 (CI: 0.39; 0.23), indicating that the percentage of visitors who stop by at least 50% of exhibit features is 0.30, that is, 30%. The values show a high degree of heterogeneity [$Q_w(100) = 1096.75$, $p < .001$; $I^2 = 93.80\%$ (CI: 96.20; 91.4)]. The estimated specific variance (τ^2) equals 1.5190 (SE = 0.2408). We did not observe any significant association with the moderating variables for dv .

—— Insert Table 3 about here ——

Inverse of Velocity (Iv)

The combined weighted estimate of the population value (Iv_\bullet) is 4.07 minutes / 100 m² (CI: 4.55; 3.59). The values show a high degree of heterogeneity [$Q_w(100) =$

8191.11, $p < .001$; $I^2 = 99.68\%$ (CI: 104.38; 94.98)]. The estimated specific variance (τ^2), equals 5.8446 (SE = 0.8509).

—— Insert Table 4 about here ——

The moderator *number of features* explains a significant proportion of the variance, although its absolute value is low. As the slope is positive, the greater the number of features the slower the speed (and the higher its inverse). Likewise, the moderator *surface area* also explains a significant proportion of the variance, although this continues to be low. In this case the slope is negative, meaning that the larger the exhibit, the more quickly visitors move through it.

Regarding *Age* variable, there is an enormous difference between categories (4.28 versus 0.70): visitors spend an average of 3.58 minutes more for every 100 m² in new exhibits than in old.

We fitted the combined model including *number of features*, *surface area*, and *age* variables, which explains 41.46% of the variance [$Q_M(3) = 51.5794$, $p < .001$]. The equation is:

$$Iv'_i = 1.41 + 0.030 \cdot N. \text{ of Features}_i - 0.464 \cdot Surface \text{ area}_i + 1.821 \cdot Age_i$$

The time/surface area relationship is greater with more features, such that with more features, visitors move more slowly. However, the time/surface area relationship is greater with greater surface areas, meaning the greater the surface area covered by visitors, the faster visitors move through the exhibit. The *age* variable adds 1.821 when

it is a new exhibit. Other models, including other moderating variables, were also explored but they do not significantly increase the predictive capacity of the model.

Stops per Feature (SF)

The combined weighted estimate of the population value (SF_{\bullet}) is 0.35 stops per feature (CI: 0.38; 0.33). The values show a high degree of heterogeneity [$Q_w(100) = 21816.10, p < .001; I^2 = 99.05\%$ (CI: 99.32; 98.78)]. The specific variance (τ^2) is equal to 0.0192 (SE = 0.0028).

—— Insert Table 5 about here ——

The *number of features* variable explains a low but significant amount of variance. The slope is negative, which means the greater the number of features the lower the mean of stops per feature made by visitors.

Assessing the asymmetry of the distributions

Conventionally, a meta-analysis includes analyses aimed at detecting possible publication bias in the field of study it refers to (Rothstein, Sutton, & Borenstein, 2005). The most studied source of bias is that toward studies with statistically significant results. But this source of bias is not relevant for our study because the original studies synthesized do not make significance tests. Furthermore, the reports have not been published in scientific journals based in peer-review processes.

—— Insert Figure 1 about here ——

Nevertheless, we have created funnel plots (Light & Pillemer, 1984) for the four ES indexes. For three of the four indicators, the values are better when the visitor sample used for the estimate is lower: *ATF* increases (they spend longer), *Iv* increases (slower movement), and *PF* increases (number of times they stop through the visit). Contrariwise, the indicator *dv* exhibits a different and more complex relationship pattern.

Discussion

The results of our analysis show that less *ATF* is spent as larger is the number of features in the exhibit; that, in the average, more time is spent in new exhibits than in old exhibits; that users move more slowly when there are more features; that in exhibits with larger surface areas the users move more rapidly; and that the time spent per feature is not related to the surface area. In general terms, our conclusions agree with those of Serrell.

However, our conclusions are reached by means of statistical tests that satisfy the assumptions of the statistical models. Furthermore, the estimates obtained with the meta-analytical procedures are more reliable, as they are unbiased, consistent, and of maximum efficiency. Using ES indices the values from different studies can be compared, as they are analyzed with a shared metric. As a consequence, our combined point estimates can be employed as a reference for evaluation and research in the field of visitors' studies. We resume them: $ATF_{\bullet} = 25.8$ seconds; $SF_{\bullet} = 0.35$ stops; $dv_{\bullet} = 30\%$ diligent visitors; $Iv_{\bullet} = 4.07$ minutes per 100 m². Analyses of complex variables are well received among researchers and professionals, as these measures allow quantifying effects and measure problems in a tangible way, offering more powerful tools for evaluation.

We have chosen the *Paying Attention* report of Serrell because it is a unique effort in the field of Museum Studies, a field where the visitors studies are confidential, not accessible to the general public. Still, our study has gone further, as we have proposed new statistical indices and have applied statistical techniques borrowed from the meta-analysis methodology. The sophistication of those techniques allows calculating efficient combined estimates and testing the role of potentially relevant moderators.

Our results show that visitor's behavior can be partially explained by factors such as the type of museum, the exhibit set-up, the characteristics of the expository features, and their layout. This gives us an idea of the type of potentially moderating variables of great importance in future meta-analysis studies. However, studying the funnel plots we have also shown unexpected strong relationships between several effect size indices and the sample size in the study. We assume that the sample size is directly related to the amount of visitors in the exhibit, so that the studies on more popular exhibits are based on larger samples. Assuming that, the funnel plots and the asymmetry tests show that when the visitor sample is small, it is also more (self) selective. As a consequence, in exhibits with smaller samples more time is spent, visitors walk more slowly, and they make more stops per feature. These effects are probably related to each visitor's specialty or personal interest. On the contrary, general population visiting massively a museum yields smaller amounts of time spent, speeded walks, and less stops.

A major contribution of our study is that the re-analysis of the database published by Serrell (1998) provides combined estimates of four indices that point trends that can serve as a reference. We know that the greater the number of elements, less time spent by visitors per item in the exhibition and fewer stops are made. Therefore, overloaded exhibitions induce superficial visits, where people do not spend enough time to the items and acquisition of new knowledge is limited. We also know that visitors devote more time to new exhibits, so that a re-design of our exhibitions can foster deeper visits. Our analysis also revealed that in the exhibitions with more elements visitors go more slowly, probably to pay attention to all the displays. However, an exhibition in which the visitor is confronted with countless pieces (usually very similar) can lead to loss of interest (and care). To design a more fulfilling and

meaningful visitor experience, it may be more useful to exhibit only those most representative and important pieces, which can focus their full attention, without overloading the visitor. Thanks to our estimates it is possible predict the average behavior of a sample of visitors, respect to the time spend per item, if we know the number of elements in our exhibition and take in account whether it is a new or an old exposure. We can also estimate the average speed at which a sample of visitors make the visit, taking in account the exhibition surface, number of items, and length of exposure. All this information can help us in planning our exposure, objectively estimating how much time the average visitor will spend in the visit and making decisions when planning the museum. The values of the four indices of effect size can be helpful for comparisons with the values obtained in evaluations of our exhibition. The goal, then, is to improve these values in new exhibitions, taking this data as references when adapting to our users.

References

- Abras, C., Maloney-Krichmar, D. & Preece, J. (2004). User-Centered Design. In Bainbridge, W (Ed.). *Berkshire Encyclopedia of Human-Computer Interaction*. Great Barrington, MA: Berkshire Publishing Group, 463–468.
- Asensio, M., & Asenjo, E. (Eds.) (2011). *Lazos de Luz Azul: Museos y Tecnologías 1, 2 and 3.0 [Blue Light Loops: Museums and 1, 2, and 3.0 Technologies]*. Barcelona: UOC.
- Asensio, M., & Pol, E. (2005). Evaluación de Exposiciones [Exhibit Evaluation]. In J. Santacana & N. Serrat (Eds.) *Museografía Didáctica [Didactic Museum Studies]*. Barcelona. Ariel, 527-630.
- Bitgood, S. (2011). *Social Designing in Museums: The psychology of visitor studies*.

Collected Essays Volumes one and two. Museumsetc.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009).

Introduction to Meta-analysis. Chichester, UK: John Wiley and sons.

<http://dx.doi.org/10.1002/9780470743386>

Borenstein, M., Hedges, L.V., Higgins, J.P.T. & Rothstein, H .R. (2010). A basic introduction to fixed-effects and random-effects models for meta-analysis.

Research Synthesis Methods, 1, 97-111. <http://dx.doi.org/10.1002/jrsm.12>

Botella, J., & Sánchez-Meca, J. (2015). *Meta-análisis en Ciencias Sociales y de la Salud [Meta-analysis of Social Science and Health]*. Madrid: Síntesis.

Botella, J., Suero, M., & Ximénez, C. (2012). *Análisis de datos en Psicología I. [Data Analysis in Psychology I]*. Madrid: Pirámide.

Carbonell, B.M. (Ed.) (2012). *Museum Studies. An anthology of contexts*. Malden, MA: Wiley-Blackwell.

Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.) (2009). *The Handbook of Research Synthesis and Meta-analysis* (2nd edition). New York: Russell Sage Foundation.

Daignault, L. (2011). *L'Evaluation Muséale. Savoirs et savoir-faire*. Québec: Presses de l'Université du Québec.

DerSimonian, & R., Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7, 177-188. [http://dx.doi.org/10.1016/0197-2456\(86\)90046-2](http://dx.doi.org/10.1016/0197-2456(86)90046-2)

Ellis, P. D. (2010). *The essential guide to effect sizes*. Nueva York: Cambridge University Press.

Falk, J. & Dierking, L.D. (2013). *The Museum Experience Revisited*. Walnut Creek, CA: Sage.

Harlow, B. (2014). *The road to results. Effectives practices for building art audiences*. N.Y.: Harlow Research and Consulting LCC.

- Heath, C. & D. Lehn, V. (2010). Interactivity and Collaboration: new forms of participation in museums, galleries and science centres. En: R. Parry (Ed.). *Museums in a Digital Age*. Abingdon: Routledge, 266-280.
- Hedges, L. V., & Olkin, I. (1985). *Statistical Methods for Meta-analysis*. Orlando, FL: Academic Press.
- Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods*, 3 (4), 486. <http://dx.doi.org/10.1037/1082-989X.3.4.486>
- Higgins, J. P., & Green, S. (Eds.). (2008). *Cochrane handbook for systematic reviews of interventions: Cochrane Book Series*. Chichester: Wiley-Blackwell. <http://dx.doi.org/10.1002/9780470712184>
- Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? *Psychological methods*, 11(2), 193-206. <http://dx.doi.org/10.1037/1082-989X.11.2.193>
- Kelley, K. & Preacher, K. J. (2012). On Effect Size. *Psychological Methods*, 17, 137–152. <http://dx.doi.org/10.1037/a0028086>
- Klingler, S., & Graft, C. (2012). In lieu of mind Reading: visitor studies and evaluation. In: Catlin-Legutko C. & Klingler, S. (Eds.) *Small Museum Toolkit*. Plymouth: AltaMira Press.
- Knell, S. J., Aronsson, P., Amundsen, A. B., Barnes, A. J., Burch, S., Carter, J., Gosselin, V., Hughes, S. A., & Kirwan, A. (Eds.). (2011). *National museums, new studies from around the world*. New York: Routledge.
- Light, R. J., & Pillemer, D. B. (1984). *Summing up. The science of reviewing research*. Cambridge, MA: Harvard University Press.
- MacDonald, S. (Ed.). (2006). *A companion to Museum Studies*. Oxford: Blackwell Pub.

<http://dx.doi.org/10.1002/9780470996836>

- Newcombe, R. G. (2012). *Confidence intervals for proportions and related measures of effect size*. Boca Raton, FL: CRC Press. <http://dx.doi.org/10.1201/b12670>
- Raudenbush, S. W. (2009). Analyzing effect sizes: Random-effects models. En H. Cooper, L.V. Hedges y J.C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2^a ed.) (pp. 295-315). Nueva York: Russell Sage Foundation.
- Serrell, B. (1998). *Paying Attention: Visitors and Museum Exhibitions*. Washington, DC: American Association of Museums.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1-48. <http://www.jstatsoft.org/v36/i03/>.
- Weil, S. (2002). *Making Museum Matter*. Washington: The Smithsonian Institution Press.

(Tabla 1)

Table 1. Summary of moderator variables.

Variable	Category	<i>N (%)</i>	<i>Mean</i>	<i>St. dev.</i>	<i>Range</i>
Topic	Science	74 (73)			
	Not Science	27 (27)			
Institution Type	Zoo or Aquarium	14 (14)			
	Natural History Museum	26 (26)			
	History/Culture Museum	19 (19)			
	Science Museum or Center	31 (31)			
	Art Museum	11 (11)			
Dioramalike	Dioramalike	82 (82)			
	Not Dioramalike	19 (19)			
Exhibit Age	Old	95 (94)			
	New	6 (6)			
Number of features			35.39	26.17	174 - 4
Surface area			389.90	259.02	1207.74 – 35.67
Sample Size			75.24	54.41	458 - 22

(Tabla 2)

Table 2. Effect size index: ATF. Summary of results for the moderator variables.

Categorical Variables							
Moderating Variable	Category	Q_b	p	ATF \bullet	(95% CI)	Q_w	p
Topic ($gl = 1$)		0.1106	.7394				
	Science ($k = 74$)			0.4407	(0.5086; 0.3729)	92.49	.0614
	Not Science ($k = 35$)			0.4184	(0.5310; 0.3058)	16.35	.9270
Institution Type ($gl = 4$)		14.103	.8424				
	Zoo or Aquarium ($k = 20$)			0.4089	(0.5379; 0.2798)	17.61	.5483
	Natural History Museum ($k = 25$)			0.5004	(0.6245; 0.3763)	29.85	.0950
	History/Culture Museum ($k = 19$)			0.4259	(0.5588; 0.2931)	8.27	.9742
	Science Museum or Center ($k = 34$)			0.4153	(0.5230; 0.3076)	44.66	.0239
	Art Museum ($k = 13$)			0.4172	(0.5933; 0.2412)	8.55	.5752
Dioramalike ($gl = 1$)		0.3138	.5727				
	Dioramalike ($k = 87$)			0.4428	(0.5073; 0.3784)	92.23	.1850
	Not Dioramalike ($k = 20$)			0.4002	(0.5336; 0.2667)	16.67	.5459
Exhibit Age ($gl = 1$)		45.938	.0321				
	Old ($k = 102$)			0.1906	(0.4206; -.0394)	0.2847	.9979
	New ($k = 6$)			0.4501	(0.5087; 0.3916)	109.04	.1375
Quantitative Variables							
Moderating Variable	Slope		(95% CI)	Z		p	R^2
N. of Features (F)	-0.0035		(-0.0014; -0.0056)	-32.830		.0010	.1026
Surface area (m ²)	0.0001		(0.0002; -0.0000)	0.8113		.04172	.0084

(Tabla 3)

Table 3. Effect size index: *dv*. Summary of results for the moderator variables.

Categorical Variables						
Moderating Variable	Category	Q_b	p	dv_{\bullet}	(95% CI)	Q_w p
Topic ($gl = 1$)	Science ($k = 74$)	0.3997	.5273	0.2858	(0.3854; 0.2109)	84.89 .1610
	Not Science ($k = 35$)			0.3444	(0.5632; 0.2106)	17.60 .8898
Institution Type ($gl = 4$)		5.3384	.2543			
	Zoo or Aquarium ($k = 20$)			0.3464	(0.6850; 0.1752)	17.02 .1983
	Natural History Museum ($k = 25$)			0.3325	(0.5416; 0.2041)	25.33 .4435
	History/Culture Museum ($k = 19$)			0.9243	(0.8030; 0.2583)	18.13 .4466
	Science Museum or Center ($k = 34$)			0.1995	(0.3190; 0.1248)	35.12 .2380
	Art Museum ($k = 13$)			0.2849	(0.6207; 0.1307)	6.40 .7799
Dioramalike ($gl = 1$)		0.1771	.6739			
	Dioramalike ($k = 87$)			0.2933	(0.3901; 0.2205)	87.18 .2994
	Not Dioramalike ($k = 20$)			0.3390	(0.6246; 0.1840)	15.63 .6180
Exhibit Age ($gl = 1$)		1.9582	.1617			
	Old ($k = 102$)			0.2885	(0.3754; 0.2216)	102.05 .2677
	New ($k = 6$)			0.6473	(1.9468; 0.2152)	0.684 .9838
Quantitative Variables						
Moderating Variable	Slope	(95% CI)		Z	p	R^2
N. of Features (F)	-0.0041	(0.0068; -0.0151)		-0.7418	.4582	.0043
Surface area (m ²)	0.0000	(0.0011; -0.0010)		0.0322	.9743	.0001

(Tabla 4)

Table 4. Effect size index: *Iv*. Summary of results for the moderator variables.

Categorical Variables								
Moderating Variable	Category	Q_b	p	Iv_{\bullet}	(95% CI)	Q_w	p	
Topic ($gl = 1$)		0.3574	.5499					
	Science ($k = 74$)			3.9871	(4.5472; 3.4270)	62.31	.8093	
	Not Science ($k = 35$)			4.3176	(5.2450; 3.3902)	50.93	.0024	
Institution Type ($gl = 4$)		2.9634	.5460					
	Zoo or Aquarium ($k = 20$)			4.4234	(5.6849; 3.1620)	11.07	.6044	
	Natural History Museum ($k = 25$)			3.3872	(4.3144; 2.4600)	25.93	.4106	
	History/Culture Museum ($k = 19$)			4.2208	(5.3071; 3.1345)	51.47	.0000	
	Science Museum or Center ($k = 34$)			4.3769	(5.2317; 3.5222)	23.48	.7946	
	Art Museum ($k = 13$)			4.1393	(5.5649; 2.7137)	2.12	.9953	
Dioramalike ($gl = 1$)		0.0023	.9618					
	Dioramalike ($k = 87$)			4.0812	(4.6134; 3.5490)	86.62	.3142	
	Not Dioramalike ($k = 20$)			4.0511	(5.1606; 2.9416)	26.82	.0823	
Exhibit Age ($gl = 1$)		14.5206	.0001					
	Old ($k = 102$)			0.7063	(-1.0742; 2.4869)	0.1327	.9997	
	New ($k = 6$)			4.2802	(4.7371; 3.8233)	116.22	.0599	
Quantitative Variables								
Moderating Variable	Slope				(95% CI)	Z	p	R^2
N. of Features (F)	0.0116				(0.0022; 0.0210)	2.4102	.0159	.0842
Surface area (m ²)	-0.3969				(-0.4869; -0.3070)	-8.6497	.0000	.1534

(Tabla 5)

Table 5. Effect size index: *SF*. Summary of results for the moderator variables.

Categorical Variables						
Moderating Variable	Category	Q_b	p	SF_{\bullet}	(95% CI)	Q_w p
Topic ($gl = 1$)		1.3938	.2378			
	Science ($k = 74$)			0.3440	(0.3759; 0.3121)	74.20 .4388
	Not Science ($k = 35$)			0.3814	(0.4347; 0.3282)	26.37 .4424
Institution Type ($gl = 4$)		9.8625	.0428			
	Zoo or Aquarium ($k = 20$)			0.3800	(0.4502; 0.3098)	16.56 .2199
	Natural History Museum ($k = 25$)			0.3727	(0.4245; 0.3209)	20.54 .7177
	History/Culture Museum ($k = 19$)			0.3698	(0.4304; 0.3092)	22.63 .2051
	Science Museum or Center ($k = 34$)			0.2941	(0.3414; 0.2469)	31.07 .4121
	Art Museum ($k = 13$)			0.4183	(0.4995; 0.3372)	9.82 .4556
Dioramalike ($gl = 1$)		0.6084	.4354			
	Dioramalike ($k = 87$)			0.3591	(0.3896; 0.3287)	88.67 .2622
	Not Dioramalike ($k = 20$)			0.3312	(0.3944; 0.2680)	11.99 .8476
Exhibit Age ($gl = 1$)		2.1787	.1399			
	Old ($k = 102$)			0.3590	(0.3871; 0.3309)	95.60 .4346
	New ($k = 6$)			0.2718	(0.3841; 0.1596)	5.02 .4124
Quantitative Variables						
Moderating Variable	Slope				(95% CI)	Z p R^2
N. of Features (F)	-0.0016				(-0.0006; -0.0026)	-3.1856 .0014 .0959
Surface area (m ²)	-0.0000				(-0.4869; -0.3070)	-0.1463 .8837 .0001

(Figura 1)

Figure 1. Funnel plot for the four ES indexes.

