RESEARCH-ARTICLE

Aesthetic Measures for Document Layouts: Operationalization and Analysis in the Context of Marketing Brochures

DAVID SCHÖLGENS, University of Cologne, Koln, Nordrhein-Westfalen, Germany

SVEN MÜLLER, University of Cologne, Koln, Nordrhein-Westfalen, Germany

CHRISTINE BAUER, University of Cologne, Koln, Nordrhein-Westfalen, Germany

ROMAN TILLY, University of Cologne, Koln, Nordrhein-Westfalen, Germany

DETLEF SCHODER, University of Cologne, Koln, Nordrhein-Westfalen, Germany

Published: 13 September 2016

Citation in BibTeX format

DocEng '16: ACM Symposium on Document Engineering 2016
September 13 - 16, 2016
Vienna, Austria

Conference Sponsors:
SIGWEB

DocEng '16: Proceedings of the 2016 ACM Symposium on Document Engineering (September 2016)
https://doi.org/10.1145/2960811.2960821
ISBN: 9781450344388

Aesthetic Measures for Document Layouts: Operationalization and Analysis in the Context of Marketing Brochures

David Schölgens, Sven Müller, Christine Bauer, Roman Tilly, Detlef Schoder
Department of Information Systems and Information Management, University of Cologne
Pohligstr. 1 – 50969 Cologne – Germany
{schoelgens; mueller; bauer; tilly; schoder}@wim.uni-koeln.de

ABSTRACT

Designing layouts that are perceived as pleasant by the viewer is no easy task: it requires a wide variety of skills, including a sense for aesthetics. When numerous documents with different content need to be created, one of the bottlenecks is to manually create appealing layouts for each document. Thus, automation for aesthetic layout creation is becoming increasingly important. Prerequisite for this automation are algorithms to measure aesthetics. While the literature proposes basic theoretical fundamentals and mathematical formulas as aesthetic measures, researchers have not operationalized these measures yet.

This paper presents the challenges associated with and the lessons learned from operationalizing 36 aesthetics measures derived from the literature for the context of marketing brochures. We measured the aesthetics of 744 brochure pages from 10 major retailers and found very strong and highly significant correlations between at least 11 of the aesthetic measures, which represent five latent aesthetic concepts. Still, most of the measures were found to be independent in our sample, and they cover a wide range of different aesthetic concepts. Nevertheless, our results suggest that retailers optimize some of these measures more than others. In terms of the aesthetic measures, retailers seem to design brochure pages in the same way regardless of which category products on this page belong to or if it is the first, last, an odd, or an even page. We propose to consider the quality values of aesthetic measures derived from our analysis of the measured brochures as target values for automated document layout creation for aesthetic marketing brochures.

Keywords

Aesthetic measures; aesthetics; document layout; layout arrangement; marketing brochures.

1. INTRODUCTION

A document has to be visually appealing and aesthetically pleasing to serve as an effective means for communication or sales [3]. Consequently, aesthetics is a key factor in document layout creation. While the various facets of aesthetics are studied in philosophy and arts [16], this paper focuses on the aspects of aesthetics that are relevant in document layouts: we consider aesthetics in terms of layouts that are perceived as pleasant by the viewer and adapt them in the context of marketing brochures.

DocEng '16, September 12 - 16, 2016, Vienna, Austria

ACM 978-1-4503-4438-8/16/09…$15.00
DOI: http://dx.doi.org/10.1145/2960811.2960821

Being able to create document layouts automatically is becoming increasingly important. A particularly interesting application domain is the creation of layouts for personalized documents in which the content is tailored to consumer needs, goals, knowledge, interests, or other characteristics [31]. Personalization leads to a high number of documents with differing content, and thus, each personalized document needs a particular layout [3, 13]. Creating different layouts for every personalized documents is time consuming and costly, particularly when done manually by designers. Consequently, automation of aesthetic layout creation is necessary, which requires automatable algorithms enabling information systems to measure aesthetics.

Within the field of document engineering, various authors have introduced and/or discussed measures for measuring the aesthetics of documents in general [2, 3, 11, 24] or specific application domains, such as newspapers [5, 6, 17], digital photo albums [4], screen design [18–20, 27], and websites [28, 30].

However, to date, an operationalization (adaption and practical implementation) of the discussed aesthetic measures is actually missing. Only Zain et al. [30] presented a small study for six measures of Ngo et al. [20], and Altaboli and Lin [1] investigated the effect of Ngo et al.'s [20] symmetry and unity on visual aesthetics in the context of one website.

We contribute to closing this research gap by adapting and implementing aesthetic measures in the context of marketing brochures. We expect high benefits from the automated layout creation of individualized documents, especially in this context. We report the challenges and lessons learned from operationalizing the provided measures for the context of marketing brochures. Furthermore, we apply our implemented solution to measure the aesthetics of 744 marketing brochure pages. This enables us to derive findings concerning the interplay of aesthetic measures and redundancies. In addition, our analysis provides interesting insights into the aesthetics of brochures for 10 major retailers in Germany. In doing so, we provide a target quality values for marketing brochures that may be particularly useful for automated document creation. Additionally, we discuss in detail the limitations of prior work, identifying future key research threads within the field of aesthetic measures.

The remainder of the paper is structured as follows: We present an overview of related work on aesthetic measures in Section 2. Section 3 describes our methodological approach for operationalizing aesthetic measures based on the literature and the measurement of marketing brochures for 10 major retailers. We report the challenges we encountered when operationalizing the aesthetic measures in our system in Section 4 and provide details on the lessons learned and our solutions. Afterwards, in Section 5, we report the results from our analysis of the marketing brochures.

After offering a critical discussion of our findings and suggesting potential future avenues of research in the field in Section 6, we conclude with a summary of the main contributions in Section 7.

2. RELATED WORK

Basically, there are three main strands of research on the concepts related to and approaches for aesthetic measurement, which go back to Ngo [18], Harrington et al. [11], and Vanderdonckt [27].

Ngo [18] named and defined 13 aesthetic measures that are relevant in the context of screen design: namely, balance, equilibrium, symmetry, sequence, cohesion, unity, proportion, simplicity, density, regularity, economy, homogeneity, and rhythm. In Ngo et al. [19] and Ngo et al. [20], these 13 measures were specified in more detail, and formulas and some empirical data were provided.

Harrington et al. [11] presented nine aesthetic measures for documents in general without a specific field of application: alignment, regularity, uniform separation, balance, white-space fraction, white-space free flow, proportion, uniformity, and page security. Balinsky [2] and Balinsky et al. [3] referred to the approach by Harrington et al. [11] and provided a more detailed discussion of symmetry, alignment, and regularity. Additionally, Riva et al. [24] built their work on Harrington et al. [11] and Balinsky et al. [3] and introduced two ways to calculate the alignment of a page.

The result of each measure described by Ngo et al. [19] and Harrington et al. [11] is a score ranging from 0 (bad) to 1 (good). Ngo et al. [19] and Harrington et al. [11] proposed a combination of the measures to compute an overall aesthetic quality score.

At almost the same time, Vanderdonckt [27] provided a list of 30 aesthetic measures for interactive applications. This work differentiated five categories: namely, physical techniques, composition techniques, association and dissociation techniques, ordering techniques, and photographic techniques. While both Ngo [18] and Harrington et al. [11] focused on the layout arrangement of a given set of objects – as we also do in our work – the work by Vanderdonckt [27] goes beyond that. First, in contrast to our approach, Vanderdonckt [27] also considered steps like object design and creation in the value chain that precede object arrangement as part of the aesthetic measurement task. For instance, Vanderdonckt [27] discussed aesthetic measures concerning photographic techniques. In our work, though, we did not consider object design and creation as variables in the process of document design; rather, we considered these issues as static, which goes in line with Ngo [18] and Harrington et al. [11]. Second, Vanderdonckt [27] took into account measures that go beyond arrangement, such as consistency, which characterizes how well visual appearance meets the subject of the placed objects. A closer look at Vanderdonckt's [27] presented aesthetic measures with respect to the arrangement of objects shows strong similarities to the aesthetic measures by Harrington et al. [11] and Ngo et al. [19] (e.g., balance, symmetry, regularity, alignment, proportion, economy, simplicity, unity). In other words, in terms of object arrangement, we consider the aesthetic measures by Harrington et al. [11] and Ngo et al. [19] as the roots.

Based on the literature on aesthetic measures, several authors have discussed application scenarios in various application domains. For instance, Chao and Lin [8] provided a template-based approach for capturing the layout of an existing page for reuse with different content. Strecker and Hennig [26] adapted some aesthetic measures to provide a flexible grid-approach for creating layouts in the context of newspapers. In addition, de Oliveira [9, 10] referred to aesthetic measures, especially homogeneity, as a requirement for an item-placement algorithm for layout creation. Lin [15] presented the concept of a "document layout design engine" but without constraints regarding the aesthetic measures of the layout. Some works [13, 23] have considered document formatting and creation as a constrained optimization problem. While the abovementioned research focused on the subject of layout and document creation, Altaboli and Lin [1], Zain et al. [29], and Purchase et al. [22] dealt with layout evaluation in the context of websites. Furthermore, a recent master thesis [14] used the aesthetic measures presented by Harrington et al. [11] and Ngo et al. [20] to investigate the influence of a few aesthetic measures on figures, including, for example, clarity and readability in the context of magazines.

Even though much of the work has engaged aesthetic measurement, especially in the context of layout and document creation, empirical validation and analysis of the interplay between the aesthetic measures are still missing [19]. We call on this research gap and present empirical validation in the context of marketing brochures and analyze the statistical relationship of the aesthetic measures in the form of their correlations.

3. METHODS

As already mentioned, Harrington et al. [11], Ngo et al. [19], and Vanderdonckt [27] introduced a set of aesthetic measures for layouts in the context of documents in general, screen design, and interactive applications. It is necessary to consolidate and discuss the suggested aesthetic measures for two reasons. First, the authors presented different numbers of aesthetic measures. This could be due to the different contexts for which the measures were suggested and the different levels of granularity; as a result, a one-to-one assignment between the aesthetic measures is impossible. Second, even though the authors sometimes use the same term, they describe different concepts, or vice versa, they use different terms for the same concept.

As Ngo et al. [19] provided the most detailed descriptions and formulas for calculating the aesthetic measures, we used this work as a starting point for explicating the theoretical foundations for our work in Section 4. Then, we analyzed Harrington et al.'s [11] aesthetic measures to complement the list of measures taken from Ngo et al. [19].

We took an iterative approach to operationalize the aesthetic measures. Challenges that occurred during the implementation informed our modification of the aesthetic measures, which we describe in Section 5.

We developed a tool in Java to calculate quality values for the 36 aesthetic measures. Brochures taken from the retailers' websites (PDF representations of printed brochures) were used as input for the aesthetic measurement tool. During the import process, double pages were automatically split into two single pages. After the import, categories and margins were manually set at the page level, whereas retailer names were set at the brochure level. Afterwards, objects – typically consisting of a headline, subheadline, image, description, and price – were marked with rectangular or polygonal shapes (see Figure 1). The aesthetic measurement is based on those shapes.

We measured 37 brochures with a total of 744 pages from the period August 2014 to September 2015 for 10 major retailers in Germany and divided them into 14 categories. Table 1 shows the distribution of the measured brochure pages across the retailers and categories. All pages of all brochures were used.

In Section 5, we provide details about our statistical analysis methodology along with the results.

Figure 1. Marked positions in measured brochure

4. CHALLENGES AND SOLUTIONS FOR THE AESTHETIC MEASURES' OPERATIONALIZATION

Ngo et al. [19], Harrington et al. [11], and Vanderdonckt [27] used different terminology in the context of layout assessment and measurement, which is partly due to their different fields of application. Like Ngo et al. [19], we use the term "object" for a content element that is placed on a page. The "layout" is the smallest bounding rectangle that wraps all the arranged objects on one page. A "page" includes the layout and the surrounding white space, as well the margins at all four edges.

Although Ngo et al. [19] and Harrington et al. [11] provided detailed theoretical foundations, we had to overcome some challenges when implementing their aesthetic measures in a software solution. First, while Ngo et al. [19] and Harrington et al. [11] provided formulas for some aesthetic measures that we could implement right away, other measures were described only theoretically. As a result, we had to create the formulas ourselves based on the theoretical foundations given (e.g., alignment, white-space free flow). Secondly, Harrington et al. [11] discussed the aesthetic measures in a generic way (without tying it to a specific application context), and Ngo et al. [19] provided the measures specifically for screen design. Using those measures in a different context – namely, for brochures – required adapting some measures to fit the brochure-specific characteristics and context (e.g., density, white-space fraction) because there seem to be differences between screen design and printed brochures.

Essentially, adapting the aesthetic measures to brochures required us to define two comprehensive settings that affect multiple aesthetic measures (Section 4.1) and eight measure-specific settings that refer to one single aesthetic measure each (Section 4.2). Eleven measures could be adapted straightaway (Section 4.3). We will provide a brief overview of the adapted measures.

In the aesthetic measurement literature provided by Harrington et al. [11] and Ngo et al. [19], several aesthetic measures consist of two or more sub-measures (e.g., balance, regularity, symmetry). To provide more detailed information and results, we worked with sub-measures and also present findings on the sub-measure level.

Table 1. Distribution of brochure pages

Category	Aldi South	Edeka	Lidl	Media Markt	Netto	Obi	Real	Rewe	Rossmann	Saturn	Σ
Car							3				3
Building Supplies	14				2	16	1				33
Office	7						1				8
Decoration	5						1				6
Electronics & Computers	12			40			28			19	99
Garden/Plants	4					8					12
Household	35	4			16	1	34	9	56		155
Domestic Appliances	2			7	4	3	8			5	29
Clothes	32				2		19	1			54
Food	39	35	24		68		64	55			285
Travel	4							2			6
Miscellaneous	4	1		1			1	6			13
Toys	3				3		5	1			12
Sports	17				1		11				29
Σ	178	40	24	48	96	28	176	74	56	24	744

Therefore, we split 10 aesthetic measures into sub-measures for a total of 36 disjunctive (sub-) measures not counting those that only represent the median of two sub-measures (an overview of all measures is shown in Table 2).

4.1 Comprehensive Settings

For the software implementation, we had to define two comprehensive settings, which refer to (1) tolerance levels and the (2) concept of dividing objects along the x- and/or y-axis.

4.1.1 Tolerance Levels

The content objects' positions were measured on a pixel basis. Thus, it would have been necessary – for instance, to get accurate alignment points – to measure the positions of the objects very precisely. As this is very costly, we used a tolerance level to simplify measurement complexity such that positions, edges, and areas are regarded as congruent.

As the limited resolution of the human eye allows for imprecision in positioning and alignment, our approach is viable. In short, for the human eye, two alignment points that are relatively close together but mathematically not exactly aligned are still perceived as aligned [3].

The introduction of tolerance levels was not only relevant in alignment calculations but also in the calculation of other aesthetic measures with respect to object positioning and size (i.e., alignment, economy, regularity, simplicity, and unity). As a tolerance level, we used 5% of the layout's width or height in pixels. For the comparison of object sizes in regard to economy and unity, we used 5% of an object's width and height.

4.1.2 Dividing Objects Along the X- and/or Y-Axis

For calculating some of the aesthetic measures (e.g., balance, symmetry), it is necessary to allocate each placed object to one quadrant or one side (i.e., left, right, top, and bottom) of the page, respectively. This is challenging when one object spans multiple quadrants. In such cases, there are basically three options for implementation: (1) dividing the object and partially attributing the object segments to multiple quadrants, (2) attributing the object to the quadrant in which the object has the most pixels, and (3) attributing the object to all quadrants in which it has pixels. We decided to use the first approach because the second and third approaches would require filling the available space in quadrants, sides, or pages more than 100%.

The first approach, which divides objects along the vertical centerline of the page and partially attributes them to the left and right sides of the page, was introduced by Harrington et al. [11] in the calculation of balance. An object that spans the left and right sides of a page is divided along the vertical centerline of the page. The left part of the object belongs to the left side of the page and the right part belongs to the right side [11].

When implementing the various aesthetic measures, we learned that division along the vertical centerline is not sufficient for some aesthetic measures (e.g., sequence). Accordingly, we augmented the concept by division along the horizontal centerline. We use this concept in balance, rhythm, sequence, and symmetry calculations.

4.2 Measure-Specific Settings

4.2.1 Alignment

Harrington et al. [11] discussed the aesthetic measure alignment. This measure's objective is to have as few alignment points as possible. A histogram-based approach for alignment calculation is proposed by Harrington et al. [11]; however, concrete formulas were not provided. The alignment point computation can be based on either left and right edges of the placed objects or their vertical centerlines. For polygonal shapes, we used the bounding rectangle for alignment calculation.

To gain more precise results, we divided alignment into four sub-measures: Alignment.top, Alignment.bottom, Alignment.left, and Alignment.right. Those sub-measures indicate how well aligned the individual edges are. The overall alignment quality is the arithmetic mean of the four sub-measures.

4.2.2 Density

Ngo et al. [19] described the concept of density in the context of screen design. It deals with how much of the available space is covered by objects. Ngo et al. [19] assumed that 50% is the optimal density level, which means that half the page is covered by objects. This assumption may be reasonable in the context of screen design, but the analysis of our data sample of marketing brochures shows that 20% white space is a more realistic condition. Our data sample has an average white space of 14.15% with a maximum of 41.93% and a minimum of 3.4%. Due to our measuring approach, it was inevitable that we marked pixels of white space between images and text blocks as content. Thus, we suggest rounding up that 14.15% to 20% to get a common formula that is appropriate if the marked white space can be captured as white space. Thus we set a density level of 80% as the optimum, which means that a layout achieves the highest possible density score if 80% of a page is covered by objects. To allow for a comparison, we implemented two alternatives: Density50 and Density80.

4.2.3 Page Security

The page security measure indicates that small objects should not be placed near the edge of a page because they seem to "fall off" the page [11]. Since this effect may be greatest at the bottom edge, Harrington et al. [11] recommended weighting the four edges differently. In addition to Harrington et al.'s [11] recommendation, our implementation normalizes the distance by half of the page's total width (on y-axis: height) to get a quality value between 0 and 1. To penalize elements that are placed closer to a problematic edge, weighting is applied in an exponential manner. For the top, left, and right edge, we used 1 as the exponential factor. For the lower edge, which is the more problematic case, we used 2 as the exponential factor.

4.2.4 Proportion

The proportion measure describes how pleasing a given height-to-width ratio is [19]. Whereas Ngo et al. [19] named five different "good" ratios – namely, square, square root of two, golden rectangle, square root of three, and double square – Harrington et al. [11] only mentioned the golden rectangle. Our implementation considers all "good" ratios named by Ngo et al. [19] and Harrington et al. [11], and all of them are considered equal. Proportion is calculated as the arithmetic mean of the two sub-measures of Proportion.object and Proportion.layout. They represent the width and height ratio of all objects respectively to the ratio of the layout.

4.2.5 Regularity

Regularity describes how evenly the content objects are distributed over a page. The regularity concept, according to Ngo et al. [19], consists of two aspects: Regularity.alignment, which focuses on minimizing the number of alignment points, and Regularity.spacing, which considers the number of different spacing distances between the arranged objects [19]. Regularity was calculated as the arithmetic mean of those two values [19].

For two reasons, we detached these aspects and divided them into two sub-measures. First, this approach allows us to provide more detailed insight into aesthetic measures and their interrelation. Second, a badly aligned layout with plenty of different alignment points could reach a relatively high regularity score if the number of different spacing points is close to 0. To consider the sub-measures as isolated measures, each sub-measure's quality value has to range from 0 to 1. Ngo et al.'s [19] formula for more than one arranged object does not handle the case of 0 spacing values, which occurs when two or more objects are arranged diagonally, without overlapping horizontal or vertical ranges. Since Ngo et al. [19] only counted vertical and horizontal spacing values, a setting where elements do not overlap horizontally and vertically would lead to 0 spacing values and subsequently to a regularity spacing quality greater than 1. For this case, we set the quality value of Regularity.spacing to 1.

While it is straightforward to determine the number of different spacing distances in a grid-based layout, for less structured layouts, it is not. Therefore, we counted all distances between two objects' edges. If one object's edge had more than one opposite edge of other objects, we counted all of these edges irrespective of their distances (see also Section 6.5).

4.2.6 Sequence

Sequence deals with how well the object arrangement facilitates eye movement, which starts in the upper-left corner and ends in the lower-right corner in the Western culture [19]. Since, according to perceptual psychologists, big objects attract the eye more than small objects, placing most of the information in the upper-left corner facilitates the reader's eye movement [19].

We implemented two versions of the sequence aesthetic measure. The first version, Sequence.w (weighted), was implemented according to Ngo et al.'s [19] approach. In this approach, the summed areas in one quadrant were multiplied by the quadrant's weight, with the upper-left quadrant's weight being 4, the upper-right quadrant's weight being 3, the lower-left quadrant's weight being 2, and the lower-right quadrant's weight being 1. We criticize that when using this approach, the upper-left quadrant has the highest value for the weighted area almost every time because the content area is multiplied by 4. Thus, it is possible that in this quadrant has less content as the other quadrants but is still seen as an optimal arrangement. Therefore, we implemented the second version, Sequence.nw (non_weighted), without multiplying the areas with the quadrants weight, and we want to compare both approaches in this paper. The weighting in the second version is still used for ordering the quadrants.

4.2.7 Symmetry

The symmetry measure describes how well objects are mirrored along an axis. This measure consists of three sub-measures, namely Symmetry.x, which compares the left and right side; Symmetry.y, which compares the top and bottom; and Symmetry.radial around a central point [19]. While Ngo et al. [19] described the calculation in detail, they did not mention their normalization approach. In our work, we normalized the x-values with the quadrants width and the y-values as well as the angle with the quadrants' height. We use height to normalize the angle only because we consider integer values (pixels) for steps in the x- and y-direction. The overall symmetry score is calculated with the arithmetic average of the three sub-measures.

4.2.8 White-Space Fraction

The aesthetic measure white-space fraction indicates that about half of a page should not be covered by objects and thus be white space [11]. While this 50% ratio may be a good target ratio in other fields of application, according to the considerations we explained in the description of density, 20% white space is a more realistic condition for our data set. Therefore, we modified Harrington et al.'s [11] aesthetic measure white-space fraction to the target ratio of 20% white space. In order to enable a direct comparison of the two possibilities of implementation, we implemented both of them. Accordingly, our Wsf50 measure has a quality value of 1 if 50% of the page is covered by objects, whereas the Wsf20 measure sees 80% of the page covered by objects as the optimum.

4.3 Adapted Measures without Modifications

4.3.1 Centered Balance

Harrington et al. [11] presented the aesthetic measure balance, which consists of left-right balance (see NgoBalance.x below) and centered balance (CBalance). CBalance describes how the visual weight of all arranged objects meets the visual center of the page, which lies slightly above the geometric center of the page [11]. CBalance is the arithmetic average of the sub-measures in the horizontal direction (CBalance.x) and vertical direction (CBalance.y).

4.3.2 NgoBalance

The aesthetic measure NgoBalance represents Ngo et al.'s [19] balance measure. NgoBalance compares the total weights of two page sides [19]. The sub-measure NgoBalance.x describes how balanced the left and the right sides of a page are. The sub-measure NgoBalance.y does the same in the vertical direction. The weight of a page side depends on the placed objects, their visual weight, and their distance from the page's geometrical center [19].

4.3.3 Cohesion

The aesthetic measure cohesion, introduced by Ngo [18] and further specified by Ngo et al. [19, 20], describes how cohesive the page is by evaluating the similarity of used aspect ratios. Therefore, the objects', layout's, and page's aspect ratios are put into relation with each other. Based on Ngo et al.'s two components in the formula for cohesion [19], we split the cohesion measure accordingly into the sub-measures Cohesion.page and Cohesion.objects. In the context of Cohesion.page, the aspect ratio of the layout is compared to the aspect ratio of the page. Within the scope of Cohesion.objects, the objects' aspect ratios are put into relation with the aspect ratio of the layout. The cohesion measure is calculated as the arithmetic mean of the two sub-measures.

4.3.4 Economy

The economy measure favors layouts with as few different object sizes (compared based on height and width) as possible to keep the layout clean and simple [19].

4.3.5 Equilibrium

The aesthetic measure equilibrium [19] requires that the center of the layout – or, more precisely, the center of mass of the arranged objects – coincides with the physical center of the page. As we did with the symmetry measure, we divided the equilibrium measure into Equilibrium.x and Equilibrium.y, splitting the horizontal and vertical part of the measure calculation. The overall equilibrium score is calculated as the arithmetic average of the two sub-measures.

4.3.6 Homogeneity

Homogeneity deals with how evenly the objects are distributed among the four quadrants of the page [19]. If all quadrants contain an equal number of objects, the homogeneity score is at its maximum [19]. To handle content that overlaps several quadrants, based on the calculation example in Ngo et al. [19], we counted this object as 1/(number of overlapped quadrants). At the end, we multiplied the number of objects on the page and within the quadrants with an integer factor of 2, 3, and/or 4 (depending on how many quadrants were overlapped) to get whole-number integers for the number of objects in the quadrants.

4.3.7 Rhythm

The rhythm measure describes how harmonic and structured a layout appears. This measure considers the variation in object arrangement horizontally (Rhythm.x) and vertically (Rhythm.y) as well as the variation in object dimensions (Rhythm.area) [19]. The overall measure is calculated with the arithmetic mean of these three sub-measures.

4.3.8 Simplicity

The simplicity measure [18] integrates two parts: (1) the number of horizontal and vertical alignment points and (2) the number of objects placed on a page.

4.3.9 Uniformity

The uniformity measure describes how well the objects' densities match the average page density, whereas variance in visual density is referred to as non-uniformity [11].

4.3.10 Unity

The aesthetic measure unity measures how well the arranged objects are perceived as "one piece" [19]. This measure was divided into two sub-measures: Unity.form describes how well the objects are related in size, and Unity.space takes into account the ratio of the white space in the layout and the white space on the page. Less white space in the layout leads to a better Unity.space quality value since the arranged objects are more closely packed together. The overall unity score is the arithmetical mean of the two sub-measures.

4.3.11 White-Space Free Flow

The white-space free flow (WsfFlow) aesthetic measure's objective is to minimize the amount of trapped white space, which cannot be reached by vertical or horizontal lines, starting from one of the four edges of the page [11].

5. RESULTS

5.1 Measure-Specific Results

In our first analysis, we examined the distribution of the aesthetic measures' quality values. We considered the quartiles that describe how 25%, 50%, 75%, and 100% of the quality values of the brochure layouts are distributed. The 50% values represent the median, indicating that half of the layouts have a higher quality value and the other half have a lower quality value. The range between 25% and 75% is called the interquartile range (IQR) and represents the size of the range for which 50% of all quality values are distributed around the median. The results are presented in Table 2.

The following measures have a low IQR ($\leq .05$) and are close to a quality value of 1 (median $\geq .92$, marked bold in Table 2): CBalance (median: .9419; IQR: .0228), CBalance.x (.9885; .0150), CBalance.y (.9208; .0328), Cohesion.page (.9791; .0306), Equilibrium (.9800; .0190), Equilibrium.x (.9884; .0150), Equilibrium.y (.9750; .0295), Proportion.layout (.9459; .0457), Rhythm.area (.9471; .0453), Sequence.w (1.0; 0.0), Wsf20 (.9919; .0121), and WsfFlow (.9979; .0073).

A high IQR value ($\geq .25$, underlined in Table 2) was found for Alignment.bottom (.2517), Alignment.top (.2507), Density80 (.3398), Homogeneity (.2591), Regularity.spacing (.2500), Sequence.nw (.5000), Unity (.2613), Unity.form (.4167), and Unity.space (.3029).

Low quality values (median $\leq .1$, underlined in Table 2) were found for Homogeneity (.0168) and PageSecurity (.0742).

5.2 Correlation Groups

We analyzed the correlation between the aesthetic measures. First, we applied the Anderson-Darling test [21] to each of the aesthetic measures to determine whether they follow normal distributions. The null hypothesis (i.e., the measure is normally distributed) was not supported for any of the measures. Hence, we assumed that all measures are not normally distributed. To account for this fact, we used non-parametric Kendall's τ [7] and Spearman's ρ [25] to assess correlations. In contrast to other measures of correlation (e.g., Pearson's r), Kendall's τ and Spearman's ρ do not require variables to be normally distributed [7, 25].

Kendall's τ is more restrictive than Spearman's ρ, generally yielding lower correlation values [7]. In our analysis, we considered only those (pairs of) metrics that were strongly correlated ($|\tau| \geq .8$ and $|\rho| \geq .8$ are generally assumed to be indicative of strong correlations [12]) with a high level of significance ($p \leq .001$) to be able to identify aesthetic measures that represent latent aesthetic concepts.

We identified five groups of aesthetic measures that correlate among each other, with each of these correlations being very strong (τ and $\rho \geq .8$) and highly significant ($p = .001$). Figure 2 represents the correlation factors of Kendall's τ (red) and Spearman's ρ (blue).

Regularity.alignment is strongly correlated with Alignment. Still, very strong correlations between the sub-measures of alignment could only be found with Spearman's ρ and, then, only between the sub-measures on the same axis.

Equilibrium and NgoBalance correlate very strongly. Their sub-measures Equilibrium.x and Equilibrium.y correlate with the corresponding NgoBalance.x and NgoBalance.y, respectively. On the x-axis, we found a strong correlation with CBalance.x; this did not apply for CBalance.y. Based on our data, we did not find any strong correlation between the sub-measures of the different axis. Only with Spearman's ρ we did find a correlation between the main measures Equilibrium and Equilibrium.y as well between NgoBalance and NgoBalance.y and between those two main measures and Rhythm.area.

Rhythm and Symmetry also correlate very strongly based on Kendall's τ. Expanding this group using Spearman's ρ, we identi-

Table 2. Quartiles of aesthetic measure quality values (bold: high median/low IQR, underlined: high IQR/low median; italic: main measures; non-italic: sub-measures)

	0%	25%	50%	75%	100%	IQR
Alignment	.2659	.5217	.6058	.6993	1.0000	.1775
Alignment.bottom	.0127	.4277	.5791	.6794	1.0000	.2517
Alignment.left	.0188	.6121	.7231	.8218	1.0000	.2097
Alignment.right	.0152	.5918	.7215	.8180	1.0000	.2262
Alignment.top	.0110	.4300	.5805	.6808	1.0000	.2507
NgoBalance	.4935	.8830	**.9219**	.9528	.9991	.0698
NgoBalance.x	.4308	.9213	**.9545**	.9787	.9999	.0574
NgoBalance.y	.2937	.8367	.9004	.9524	1.0000	.1157
CBalance	.7992	.9294	**.9419**	.9523	.9960	**.0228**
CBalance.x	.8288	.9797	**.9885**	.9947	1.0000	**.0150**
CBalance.y	.7530	.9024	**.9208**	.9352	.9998	**.0328**
Cohesion	.5382	.7465	.7920	.8301	.9992	.0837
Cohesion.objects	.1871	.5136	.6076	.6905	1.0000	.1769
Cohesion.page	.7511	.9604	**.9791**	.9910	1.0000	**.0306**
Density50	.0669	.2040	.2599	.3449	.8386	.1409
Density80	.1672	.5101	.6497	.8499	.9999	.3398
Economy	.0400	.1250	.2000	.2500	1.0000	.1250
Equilibrium	.8446	.9694	**.9800**	.9883	.9995	**.0190**
Equilibrium.x	.8286	.9797	**.9884**	.9947	1.0000	**.0150**
Equilibrium.y	.7707	.9587	**.9750**	.9882	1.0000	**.0295**
Homogeneity	.0000	.0000	.0168	.2591	1.0000	.2591
PageSecurity	.0014	.0419	.0742	.1357	.9926	.0938
Proportion	.6333	.8733	.8992	.9243	.9965	.0511
Proportion.layout	.7444	.9243	**.9459**	.9701	.9999	**.0457**
Proportion.object	.3924	.7940	.8559	.9034	.9965	.1094
Regularity	.1667	.5357	.6250	.7045	1.0000	.1688
Regularity.alignment	.0000	.5000	.6000	.6875	1.0000	.1875
Regularity.spacing	.0833	.5000	.6429	.7500	1.0000	.2500
Rhythm	.6055	.7512	.7934	.8501	.9962	.0988
Rhythm.area	.6585	.9207	**.9471**	.9660	.9967	**.0453**
Rhythm.x	.4740	.6745	.7462	.8225	1.0000	.1481
Rhythm.y	.4475	.6205	.6850	.7927	.9967	.1723
Sequence.nw	.0000	.0000	.2500	.5000	1.0000	.5000
Sequence.w	.2500	1.0000	**1.0000**	1.0000	1.0000	**.0000**
Simplicity	.0566	.1500	.2000	.3000	1.0000	.1500
Symmetry	.6020	.7421	.7895	.8406	.9974	.0984
Symmetry.radial	.4275	.6852	.7574	.8270	.9963	.1419
Symmetry.x	.4865	.7676	.8380	.9191	.9996	.1515
Symmetry.y	.4275	.7248	.7960	.8662	.9974	.1414
Uniformity	.5942	.8420	.8847	.9145	1.0000	.0725
Unity	.0844	.3589	.4890	.6202	1.0000	.2613
Unity.form	.0667	.2500	.4286	.6667	1.0000	.4167
Unity.space	.0000	.3761	.5130	.6791	1.0000	.3029
Wsf20	.9266	.9851	**.9919**	.9972	1.0000	**.0121**
Wsf50	.1293	.3664	.4522	.5708	.9739	.2044
WsfFlow	.9172	.9927	**.9979**	1.0000	1.0000	**.0073**

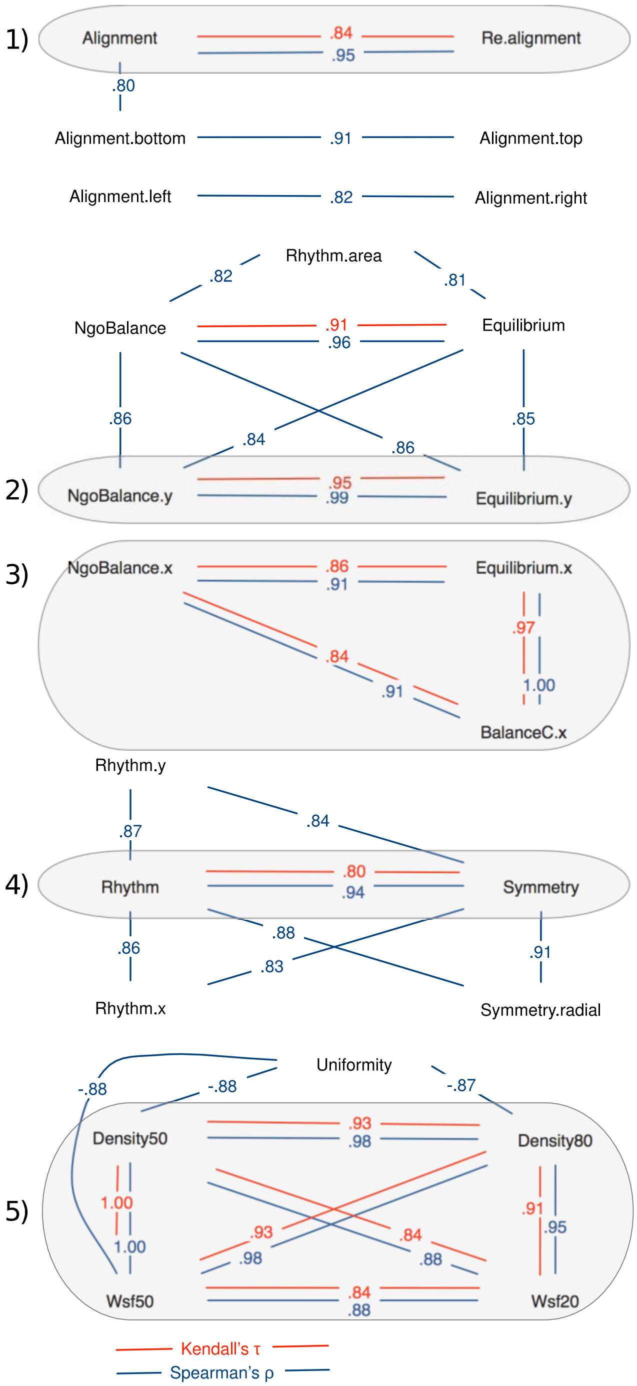

Figure 2. Correlations by Kendall's τ (red) and Spearman's ρ (blue); marked groups of measure concepts

fied correlations with respect to the sub-measures Rhythm.x and Rhythm.y as well Symmetry.radial. Very strong correlations to the other sub-measures Symmetry.x, Symmetry.y, and Rhythm.area were not found.

The largest group of inter-correlating aesthetic measures consists of Density50, Density80, Wsf20, and Wsf50. All four measures correlate very strongly with each other. Using Spearman's ρ, Uniformity also correlates with three of these measures (i.e., Density50, Density80, and Wsf50).

5.3 Retailer, Category, and Page Type

To investigate whether the quality values of the aesthetic measures differ across retailers, we compared the quality values for the pages of each retailer. From Table 1, we see that retailers represent only a few product categories. Accordingly, to exclude effects from the product categories' characteristics, in a second step, we compared the retailers by only considering pages of a particular product category. Thus, we compared retailers by using only pages of the three largest product categories separately: electronics and computers, food, and household.

In all three categories, we found that the retailer Aldi differs from the other retailers. Interestingly, its IQR is frequently larger and/or its median is shifted from those of the other retailers. This especially applies for Alignment.bottom, Alignment.left, Alignment.right, Alignment.top, Economy, PageSecurity, Regularity.spacing, Simplicity, Uniformity, and Unity.space.

Compared to the other retailers, Rewe and Rossmann have better quality values in the household category for Density50, Density80, and Wsf50 and worse quality values for Uniformity (cf. negative correlation as presented in Section 5.2). Rossmann has worse quality values for Simplicity in the household category.

In the food and household categories, Edeka has lower quality values in comparison to the other retailers in the measures Economy, PageSecurity, Regularity.spacing, Unity.form, and Unity.space. Edeka's brochures did not contain pages in electronics and computers.

Comparing the aesthetic measures, no significant differences were found for the different product categories or between first, last, odd, or even pages.

6. DISCUSSION

From our results, several conclusions can be drawn with respect to the single measures, relationships between the measures, and the design specifics of retailers. Furthermore, the implications of our modifications during the operationalization are discussed. We point out that our findings may be limited to the scope of marketing brochures and further characteristics of our sample data, such as evaluated retailers, categories, etc.

6.1 Discussion Regarding Single Measures

We found that some measures have a notably small IQR and that some have a notably large IQR. Those measures with a small IQR are interestingly also close to a quality value of 1. One explanation for the small IQR may be that designers have created the layouts by explicitly optimizing for this measure. It may also be "easy" to achieve a high quality value for these dimensions. Another explanation could be that the respective measure concepts are inappropriate to distinguish layouts properly.

High IQR of a measure may be a result of the difficulty to comply with the measure (e.g., because a measure contradicts another one) or designers' unfamiliarity with or ignorance of the respective aesthetic concept. Furthermore, a measure's concept could also be misunderstood or realized in a wrong way, which would thus lead to a measure with less explanatory power.

Except the measures with high variance (i.e., Alignment.bottom, Alignment.top, Density80, Homogeneity, Regularity.spacing, Sequence.nw, Unity, Unity.form, and Unity.space), every other measure leads to similar quality values for all 744 layouts. Thus, we surmise that the designers of the analyzed retailer brochures optimized the layouts in the same way regarding the evaluated measures.

Given the low quality values (close to 0), it seems that the measures PageSecurity and Simplicity are not so important for designers. An alternative explanation could be that these measures contradict other measures. However, our correlation analysis did not identify any negative correlations for these two measures.

6.2 Discussion of Correlations

Uniformity was the only measure identified to have strong negative correlations with the other measures. However, the negative correlation of Uniformity with Density50, Density80, and Wsf50 reflects the fact that designers have to balance an optimal Uniformity quality level and bad quality values of the latter measures or vice versa.

Furthermore, we found groups of measures that have strong correlations between each other. We assume that measures that make up such groups all represent the same aesthetic concept, whereas a different group of measures represents a different aesthetic concept.

The very strong correlations of the measures Density50, Density80, Wsf20, and Wsf50 (measure Concept 5 in Figure 2) are intuitive: density measures the amount of content, whereas white-space fraction measures the amount of white space, which is the inverse and can be derived from the density. In fact, these four measures only differ in the target value for the amount of content, which we will further discuss in Section 6.3. Based on this, we propose to use either density or white-space fraction.

In the same correlation group (i.e., Density50, Density80, Wsf20, and Wsf50), we also presented Uniformity (cf. Figure 2). Using Spearman's ρ for computing the correlation factors, Uniformity correlates with all measures in the group except Wsf20. Examining the scatter plots revealed missing examples for the following cases in our sample: only a few layouts in our data sample have a density close to 50% or 100% (max: 96.6%; min: 58.07%), thus leading to no measured quality values for uniformity close to 0 and quality values of 1 for Density50 and Wsf50. Taking such layouts into account, the gap in the scatter plot could be filled and the correlation found would probably be strong. Another factor that has to be taken into account is that we operationalize uniformity by calculating the visual weight based on a content object's area. We suggest investigating further on this measure integrating luminance, color, etc. to calculate the visual weight. With the current operationalization and based on our findings, we do not suggest that uniformity belongs to the same aesthetic concept (measure Concept 5) as density and white-space fraction, but it might be conceivable when considering also luminance, color etc. for the visual weight.

Furthermore, as their names already suggest, we found a very strong correlation between Alignment and Regularity.alignment (measure Concept 1 in Figure 2). In fact, the computation of these measures is very similar. Against this background, we propose to substitute the calculation of Regularity.alignment by the Alignment concept that considers all four edges as described in Section 4.2.

The next group of strongly correlating measures that we found is the one between NgoBalance and Equilibrium, NgoBalance.x and Equilibrium.x, as well as NgoBalance.y and Equilibrium.y. We find clear similarities between the formulas of NgoBalance and Equilibrium and their sub-measures. Where NgoBalance uses Euclidian metrics to calculate the distance between object and center of a page, Equilibrium analyses the x- and the y-axis separately; besides this difference, these measures are calculated the same way. Due to the similar formulas and the strong and significant correlations, we propose to consider them as one measure concept for each axis or for the main measure (measure Concepts 2 & 3 in Figure 2). Interestingly, we could not identify a strong correlation between NgoBalance and its sub-measures; the same applies for Equilibrium. We conjecture that one sub-measure balances the quality value of another one.

For the x-axis, we could identify a strong correlation of NgoBalance.x and Equilibrium.x with CBalance.x. The formula of CBalance sums up all center positions on the x-axis and weights them before they are related to the center of the page; this procedure is very similar to the one of NgoBalance and Equilibrium. Therefore, we suggest subsuming CBalance.x under the same concept as NgoBalance.x and Equilibrium.x (measure Concept 3 in Figure 2). In contrast to the x-axis, only weak correlations were found concerning CBalance.y (CBalance.y/NgoBalance.y: $\tau = -.36$, $p = .001$; CBalance.y/Equilibrium.y: $\tau = -.38$, $p = .001$). Probably this could result from the fact that the visual center lies slightly above the geometric center of the page. Here further investigations have to be done, e.g. with another data set.

The strong correlation found between Rhythm and Symmetry form the fourth group. The formulas of Rhythm take values for x, y, and area into account; symmetry uses in addition to x and y values for width, height, angle, and distance. This might be reason why their sub-measures do not correlate but their combination in the main measure correlates (measure Concept 4 in Figure 2).

6.3 Discussion of Page Specifics

Interestingly, we found that the brochures of Aldi are distinct in some measures (cf. Section 5.3) compared to the other retailers' brochures. Delving into details reveals that Aldi uses a lot more full-page images in their brochures than other retailers do. In many cases these images span over a double page while products are only placed on one of the two pages. Furthermore, while Aldi places on average only 3.44 products on a page, overall brochures place 8.59 products on a page (with a maximum of 23.61 products on Rossmann pages). This may explain the differences in the measures.

Furthermore, one cause for the lower quality values of Edeka brochures for Economy, Homogeneity, PageSecurity, Regularity, Simplicity, and Unity may arise from the fact that Edeka arranges products in an entangled way in its brochures, not following a grid-based approach. Additionally, with 13.98 products on average, Edeka presents a relatively high amount of products per page. In fact, those pages may seem complex, overloaded, and confusing, which would explain the rather low quality values.

Interestingly, our data sample does not suggest any significant differences in the layouts of different page types. We assumed that we would find different quality values, for instance, for the first or the last page because these are more prominent and try to attract customers' attention. However, this was not the case given the measures under investigation.

Furthermore, it appears that designers use the aesthetic concepts in the same way over all brochure pages irrespective of the page type or the product category presented. Against this background, we propose: In order to automatically generate layouts that are aesthetically perceived as similar to those we evaluated, the aesthetic measures' quality values of the measured brochures can be considered as target values for the aesthetic layout creation of marketing brochures.

6.4 Discussion of Our Operationalization

With regard to the four possible directions for alignment, the quality values for the alignment of left and right edges was slightly better than for the top and bottom edges. Still, we could not find any evidence that designers would prioritize top and left edges. If products are sized the same in a grid-based approach, then the left and right and top and bottom edges will align naturally. In such a case, it would be sufficient to evaluate only one direction for both the vertical and horizontal axes. Still, for cultural reasons, we suggest further investigating all four directions because, for instance, in the Western culture, the left edge may be more important than the right edge, whereas the opposite is true in Arab cultures.

Furthermore, from our data set, we see that retailers optimize for a white space of about 20% and content density of about 80%. Accordingly, for our data sample, it was useful to adapt the target value for white space to 20% and content to 80% for the density measure to achieve better quality values. Still, the level may depend on the context for which the layouts are created. In addition, further in-depth research is necessary to determine consumers' aesthetic preferences.

Comparing our two approaches for page security, we found that the version that corresponds to Ngo et al. [19] leads to a quality value of 1 for almost all layouts. As a result, this measure can hardly be used to distinguish different layouts. The approach we proposed leads to quality values ranging from 0 to 1, with most layouts achieving values from 0 to 0.5. Although this approach yields worse overall quality values, it is better suited to differentiate between the layouts. Therefore, we suggest using our approach or a third "intermediate" approach that uses less heavy weights.

6.5 Future Avenues of Research

For the measures centered balance and uniformity, Harrington et al. [11] proposed using visual weights. In our study, we operationalized the content's area as a factor for the visual weight. Thus, a text block has the same visual weight as a dark image. We suggest exploring color, luminance, etc., as additional dimensions for calculating visual weight in a more detailed study. We expect interesting insights from such a study, whereas we believe that our approximation does not have large impact on the results, and if so, then the visual weights would only affect those two measures. Additionally, we suggest introducing visual weight to conceptualize density, NgoBalance, white-space fraction, and uniformity instead of using only the area concept because a text's density may appear less than that of an image. For instance, Ngo et al. [19] proposed this for NgoBalance in their future research.

The calculation of different spacing sizes (n_spacing) in regularity was described in Ngo et al. [19] only for the horizontal and vertical directions. As described, we measured the content shapes as polygons. Using our approach, we can additionally have diagonal distances, for example, between a ski and a ski pole that are positioned parallel in a diagonal direction. What size has to be used? If they were not positioned parallel but convex, which size has to be used for the spacing? This aspect needs to be discussed in more detail. In our study, we restricted the measurement to horizontal and vertical distances by using the bounding rectangles for polygons as an approximation. Another aspect for the calculation of n_sizes is the case in which an object has a connection to another object through a small gap between two other elements (e.g., the header of a page has a small connection to the footer through the gap between the two columns). In our paper, we counted every connection regardless of the size of the gap. However, using a minimum size threshold for the gap is also a conceivable alternative. This would probably match with the viewer's visual perception, but this notion needs to be researched.

In our study, we treated all pages as single pages regardless of whether they were used in a double-page design. Some of the marketing brochures used one background image for the whole double page, thus, connecting those pages. In the context of double pages, it would be interesting to apply the measures to a double page instead of examining both pages independently. This becomes especially necessary if content objects overlap both pages. Measures that would be particularly affected in this context are page security, symmetry, and NgoBalance. For page security, the left edge of a right page would have to be weighted differently than the left edge of a left page.

Our weighting of the edges for page security seems to be a good starting point. Still, future research could also investigate the effects of cultural aspects regarding the aesthetic measures. Cultural characteristics may impact alignment (cf. Section 6.1), CBalance, page security, and sequence. The optical center in CBalance may need to be adapted, for example, in cultures where scripts are written from bottom to top. Furthermore, for the measure page security, reading direction may influence the weights needed for the different page edges. In cultures that read from right to left, it may be preferable to place most information in the upper-right corner and adjust the sequence of the quadrants. Thus, sequence would have to be adapted.

7. CONCLUSION

Automating the creation of aesthetic layouts is an increasingly important topic. First, however, automatable algorithms that enable information systems to measure aesthetics are required. While the literature has proposed basic theoretical fundamentals and mathematic formulas for aesthetic measures, an operationalization of those measures and empirical validation had not been published yet.

Against this background, the present paper provided four main contributions. First, we adapted and implemented the aesthetic measures in the context of marketing brochures. We reported the lessons learned and also provided solutions for how to adapt and modify the fundamentals and formulas so they can be meaningfully applied in practice.

Second, we measured a set of 744 marketing brochure pages for 10 retailers along the 36 aesthetic measures as described in the literature. The results suggest that the measures are applicable and the analysis provides interesting insights into designers' operationalizations of the aesthetic concepts in practice. For instance, we identified that there are some measures for which all retailers' brochures achieve rather high quality measures, some for which all retailers have low quality measures, and only nine measures with high IQR. As the achieved quality values per aesthetic measure show a rather coherent picture among the analyzed retailers, we propose to consider this "aesthetic profile" as target values for aesthetic layout creation for marketing brochures. This target values may be particularly useful for automated document creation.

Third, we presented five groups of measures that have very strong correlations between each other and may thus be consolidated as each group appears to represent the same aesthetic concept.

Fourth, we outlined the limitations of prior work. To overcome these limitations, we discussed future key research threads within the field of aesthetic measures. As aesthetics is a key factor in automated document layout creation, aesthetic measures and their operationalizations are a challenging but important topic. We need to address these measures to be able to provide automatable algorithms that can measure and create layouts that are aesthetic. Only then will we be in a position to automatically design document layouts that are perceived as visually appealing and aesthetically pleasing by the viewer.

8. REFERENCES

[1] Altaboli, A. and Lin, Y. 2012. Effects of Unity of Form and Symmetry on Visual Aesthetics of Website Interface Design. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 56, 1, 728–732.

[2] Balinsky, H. Y. 2006. Evaluating Interface Aesthetics: Measure of Symmetry. *Proc. SPIE* 6076.

[3] Balinsky, H. Y., Wiley, A. J., and Roberts, M. C. 2009. Aesthetic Measure of Alignment and Regularity. In *DocEng '09. Proceedings of the 2009 ACM Symposium on Document Engineering*. ACM, New York, NY, 56–65. DOI=10.1145/1600193.1600207.

[4] Bergmann, F. B., Manssour, I. H., Silveira, M. S., and de Oliveira, J. B. S. 2013. Automatic Layout Generation for Digital Photo Albums: A User Study. In *Proceedings of the 15th International Conference on Human-Computer Interaction: Users and Contexts of Use - Volume Part III*. HCI'13. Springer-Verlag, Berlin, Heidelberg, 117–126. DOI=10.1007/978-3-642-39265-8_13.

[5] Buhr, M. 1995. *Aesthetics of Newspaper Layout - and a Survey on Architecture Determining Algorithms,* Computer Science Department, Aarhus University, Denmark.

[6] Buhr, M. 1996. *Newspaper Layout Aesthetics Judged by Artificial Neural Networks*.

[7] Capéraà, P. and Genest, C. 1993. Spearman's ρ is larger than kendall's τ for positively dependent random variables. *Journal of Nonparametric Statistics* 2, 2, 183–194.

[8] Chao, H. and Lin, X. 2005. Capturing the Layout of Electronic Documents for Reuse in Variable Data Printing. In *Proceedings of the 2005 Eight International Conference on Document Analysis and Recognition (ICDAR'05)*, 940. DOI=10.1109/ICDAR.2005.68.

[9] de Oliveira, J. B. S. 2008. Two Algorithms for Automatic Document Page Layout. In *DocEng '08. Proceedings of the eighth ACM Symposium on Document Engineering*. ACM, New York, NY, 141–149. DOI=10.1145/1410140.1410170.

[10] de Oliveira, J. B. S. 2009. Two Algorithms for Automatic Page Layout and Possible Applications. *Multimedia Tools and Applications* 43, 3, 275–301.

[11] Harrington, S. J., Naveda, J. F., Jones, R. P., Roetling, P., and Thakkar, N. 2004. Aesthetic Measures for Automated Document Layout. In *DocEng '04. Proceedings of the 2004 ACM Symposium on Document Engineering*. ACM, New York, NY, 109–111. DOI=10.1145/1030397.1030419.

[12] Henn, M., Weinstein, M., and Foard, N. 2009. *A critical introduction to social research*. SAGE, Los Angeles.

[13] Hurst, N., Li, W., and Marriott, K. 2009. Review of Automatic Document Formatting. In *DocEng '09. Proceedings of the 2009 ACM Symposium on Document Engineering*. ACM, New York, NY, 99–108. DOI=10.1145/1600193.1600217.

[14] Kivelä, I.-M. 2012. Aesthetic Measures for Automated Magazine Layout on Tablet Devices. *Aalto University School of Science Department of Media Technology: Master thesis*.

[15] Lin, X. 2006. Active Layout Engine: Algorithms and Applications in Variable Data Printing. *Computer-Aided Design* 38, 5, 444–456.

[16] McWhinnie, H. J. 1968. A review of research on aesthetic measure. *Acta Psychologica* 28, 363–375.

[17] Meissner, M. 2007. *Zeitungsgestaltung*. Journalistische Praxis. Econ, Berlin.

[18] Ngo, D. C. L. 2001. Measuring the Aesthetic Elements of Screen Designs. *Displays* 22, 3, 73–78.

[19] Ngo, D. C. L., Teo, L. S., and Byrne, J. G. 2002. Evaluating Interface Esthetics. *Knowledge and Information Systems* 4, 1, 46–79.

[20] Ngo, D. C. L., Teo, L. S., and Byrne, J. G. 2003. Modelling Interface Aesthetics. *Information Sciences* 152, 25–46.

[21] NIST/SEMATECH. 2016. *e-Handbook of Statistical Methods: 1.3.5.14. Anderson-Darling Test*. http://www.itl.nist.gov/div898/handbook/eda/section3/eda35e.htm. Accessed 21 March 2016.

[22] Purchase, H. C., Hamer, J., Jamieson, A., and Ryan, O. 2011. Investigating Objective Measures of Web Page Aesthetics and Usability. In *Proceedings of the Twelfth Australasian User Interface Conference*. AUIC 2011 117. Australian Computer Society, Darlinghurst, Australia, 19–28.

[23] Purvis, L., Harrington, S. J., O'Sullivan, B., and Freuder, E. C. 2003. Creating Personalized Documents. In *DocEng '03. Proceedings of the 2003 ACM Symposium on Document Engineering*. ACM, New York, NY, 68. DOI=10.1145/958220.958234.

[24] Riva, A. D., Seki, A. K., de Oliveira, J. B. S., and Manssour, I. H. 2010. Two New Aesthetic Measures for Item Alignment. In *DocEng '10. Proceedings of the 2010 ACM Symposium on Document Engineering*. ACM, New York, NY, USA, 263–266. DOI=10.1145/1860559.1860619.

[25] Salkind, N. J. and Rasmussen, K., Eds. 2007. *Encyclopedia of Measurement and Statistics*. SAGE Publications, Thousand Oaks, Calif.

[26] Strecker, T. and Hennig, L. 2009. Automatic Layouting of Personalized Newspaper Pages. In *Operations Research Proceedings 2008*, B. Fleischmann, K.-H. Borgwardt, R. Klein and A. Tuma, Eds. Springer, 469–474. DOI=10.1007/978-3-642-00142-0_76.

[27] Vanderdonckt, J. 2003. Visual Design Methods in Interactive Applications. *Chapter* 7, 187–203.

[28] Vasilyeva, O. and Lin, A. C. H. 2011. A Theoretical Framework For Understanding Aesthetic Experiences In Relation To Website Design And Utilitarian Outcomes. In *PACIS 2011 Proceedings*. Queensland University of Technology, Brisbane, Australia.

[29] Zain, J. M., Tey, M., and Goh, Y. 2011. Probing a Self-Developed Aesthetics Measurement Application (SDA) in Measuring Aesthetics of Mandarin Learning Web Page Interfaces. *CoRR* abs/1101.1606, 31–40.

[30] Zain, J. M., Tey, M., and Soon, G. Y. 2008. Using Aesthetic Measurement Application (AMA) to Measure Aesthetics of Web Page Interfaces. In *Fourth International Conference on Natural Computation, 2008. ICNC '08*. IEEE, Jinan, 96–100. DOI=10.1109/ICNC.2008.764.

[31] Zimmermann, A., Specht, M., and Lorenz, A. 2005. Personalization and Context Management. *User Modeling and User-Adapted Interaction* 15, 3-4, 275–302.