An Evaluation of Paper Brightness Indices

Bruce Myers, Hanyi Cheng[1]

[1]Rochester Institute of Technology

Keywords: paper brightness, TAPPI brightness, spectrophotometers, variance, optical brightening agents (OBAs), paper manufacture

Abstract

The primary goal of the present research is to establish the variance that can be expected when measuring TAPPI brightness as a traceable metric with a laboratory-grade meter compared to brightness indices measured by production-grade spectrophotometers. It is also recognized that the variance between the analyzed metrics may differ based upon the type of paper measured. Therefore, this research also aims to examine the brightness across several types of papers used in the printing and packaging industry. Also, an important property of many paper substrates known as optical brightness agents (OBAs) is analyzed with a goal of ascertaining its influence on brightness readings among the evaluated instruments and metrics. OBAs are commonly used chemical materials in paper pulp, which can enhance the "brightening" effect on the appearance of the material. After data collection and analysis, it is concluded that the brightness indices from the production-grade instruments used in this study result in higher readings than TAPPI brightness readings. The level of OBAs in the paper is also found to influence paper brightness readings.

Introduction

The Technical Association of Pulp and Paper Industry (TAPPI) defines peer-reviewed standards to ensure that paper and paperboard products meet industry best practices. These include a traceable metric known as TAPPI brightness (T452), designed to measure paper brightness and provide a means to compare and evaluate papers. A function of the reflectance characteristics of paper, TAPPI brightness is measured at CIE Standard 45°/0° geometry at 457-nanometer reflectance (IPSTESTING, 2017). Strict adherence to TAPPI brightness mandates a laboratory-grade measurement instrument that is dedicated mainly for the purpose of reading the metric. In contrast, production-grade spectrophotometers, which are commonly utilized by printers for quality assurance and process control, frequently provide a brightness index. Manufacturers of these measurement instruments contend that these brightness indices, while not officially meeting the requirements of TAPPI brightness, nonetheless provide a useful metric that, at a minimum, can be utilized in a relative manner by practitioners. The present research seeks to compare TAPPI brightness as measured by a laboratory-grade meter to the brightness indices of seven different production-grade spectrophotometers. The variance of the resultant measurements is analyzed across eleven different paper samples representative of a range of commonly-used printing substrates.

The prominence of paper brightness in the labeling of paper suggests the importance of this particular metric to print buyers. With production-grade color measurement instrumentation offering brightness indices, it is potentially important for stakeholders to understand differences between these indices and TAPPI brightness. A search of ProQuest for Dissertations and Theses, Scholarly Journals, Trade Journals, and Conference Papers and Proceedings, with the terms "TAPPI Brightness" or "TECHNICAL ASSOCIATION OF THE PULP PAPER INDUSTRY" in the abstract yielded no studies that examined TAPPI Brightness versus other brightness indices. Further, a search of the website for The Journal of Print and Media Technology Research (http://iarigai.com/publications/journals/) published by the International Association of Research Institutes for the Graphic Arts Industry (IARIGAI) from 2012 through March of 2018 yielded no articles that investigated TAPPI Brightness versus brightness indices. Likewise, a search of the website for the International Circular of Graphic Education and Research (https://www.internationalcircle.net/circular/issues/08_01/) published by The International Circle of Educational Institutes for Graphic Arts Technology and Management from 2008 through 2018 did not result in any articles evaluating TAPPI brightness versus other brightness indices. It is therefore suggested that the present study could represent a contribution to the literature.

Research Questions

There are three research questions examined in the present study. Each question includes a sub-question to evaluate the variance between paper samples with differing OBA levels:

1. How much variance can be expected when reading TAPPI Brightness using laboratory-grade Technidyne Tappi Brightness Meter?
   a. Is there a difference in TAPPI brightness variance when measuring paper with High OBA versus papers Low or No OBA?
2. How much variance can be expected when reading "brightness indices" with production-grade instruments?
   a. Is there a difference in the variance of brightness indices when measured with production-grade instrumentation on papers with different OBA characteristics?
3. How close are the brightness indices as measured with production-grade instruments to TAPPI brightness as measured with a laboratory-grade instrument?
   a. Is there a difference in how close brightness indices as measured with production-grade instruments are to TAPPI brightness as mea sured with a laboratory-grade instrument when measuring papers with different OBA character istics?

Methodology

An experiment was designed to evaluate the readings of TAPPI brightness and other brightness indices. After preparing samples and calibrating the instruments, data were collected and analyzed. Different levels of OBAs were also considered as a variable during the analysis to ascertain any possible effect between the different groups of paper samples.

Paper Samples

The researchers chose eleven paper samples of various types from different companies and mills. Although all paper samples utilized would be described as white, these samples varied in color cast, surface characteristics, and OBA level: the goal was to incorporate a range of papers with differing levels of OBAs ranging from none to high, with which to evaluate the variance in brightness readings across the various instrumentation.

Before their paper brightness indices were collected, paper samples were separated into groups based on their OBA level. The "OBA-Check" function on Techkon SpectroDens was used to measure the influence of OBAs in the samples, and to categorize them accordingly. Two papers exhibited no OBA, and, according to the table from Techkon SpectroDens Manual, five were categorized as low to moderate OBA while the remaining four types of paper were determined to be high OBA, as illustrated in Table 1.

Table 1: OBA-Check results

Types of Paper Sample	OBA Categorization	OBA Amount
Iggesund Invercote T	no	no
SBS	no	no
Verso Sterling Litho C1S	moderate	3.8
NewPage Ecopoint	moderate	4.1
5516 Verso Oxford	moderate	4.1
Mosaic	moderate	4.5
5516 Sappi LusterCote	moderate	4.6
5516 Evergreen	high	6.5
Iggesund Invercote G	high	8.1
Verso Sterling Premium Dull	high	8.7
Neenah	high	11.2

Instruments

Six different measurement instruments were used in the analysis. For the traceable TAPPI brightness index, the Technidyne Brightimeter S-4 at the Rochester Institute of Technology (RIT) Printing Applications Lab (PAL) was utilized, as this instrument measures TAPPI Brightness and conforms to TAPPI Official Test Method T452. Five different production-grade spectrophotometers were selected to be used for other brightness indices as shown in Table 2. All instruments were certified by their respective manufacturer within 12 months prior to the readings.

Table 2: Production-grade instruments

Manufacturer	Model	Status
Konica Minolta	FD-7	Current
X-Rite	eXact	Current
X-Rite	530	Legacy
X-Rite	939	Legacy
X-Rite	SpectroEye	Legacy

Limitations

Limitations in the present study can be thought of as centered on two primary areas: the materials and equipment utilized, and the methods employed. Each is subsequently reviewed.

Limitations in materials and equipment utilized

The study is limited to color measurement instruments that are unidirectional in nature in terms of instrument geometry: the benchmark laboratory-grade Technidyne Brightimeter S-4 is 45°/0° geometry, and each of the production-grade instruments utilized is either 45°/0° or 0°/45°, therefore the present study does not analyze instruments with other geometries (e.g.: spherical d/8°). Further, the study included both instrument models that are no longer sold as new (known as "legacy" instruments) and models that are currently sold, as presented in Table 2.

All of these instruments are commonly used in printing production facilities. While it is notable that another popular model, the Techkon SpectroDens, was not utilized in the analysis of instruments, it is important to recognize that this particular model does not offer a Brightness Index per se. This device does, however offer an "OBA Check" function, which measures the amount of optical brightening agents present in substrates, and ranks the amount of OBA in an ordinal scale: as previously stated this scale was used to categorize the paper samples used in the study.
The decision to only categorize paper by OBA presence andnot by other physical and optical characteristics, such as gloss, texture, and cast, can also be viewed as a limitation to the present analysis.

In addition to the measurement instruments selected, the paper samples used in the present study represent a range of printing papers, from "No" OBA to "High" OBA. While representing a range, the papers selected are not an exhaustive representation based on a statistical analysis of paper popularity.

Limitations in methods utilized

The present study did not conduct full gage repeatability and reproducibility studies (gage R&R) of the measurement instruments used, as suggested by Cepova, Kovacikoya, Cep, Klaput and Mizera (2018) and Sloop (2009). Such gage R&R studies trace three sources of variation in manufacturing processes: variation due to the product itself, variation as a result of the operators, and variation caused by the equipment. The methods utilized here involved a single operator taking readings over several days, and a paired-samples t test suggested the amount of variation that can be expected in the respective measurement instrument.
Further, the method of ranking papers by OBA presence as defined by Techkon SpectroDens can also be viewed as a limitation of the present study. As indicated, this device quantifies the amount of OBA and ranks measurements on a ordinal scale. While offering a convenient method of categorizing paper, other methods of categorizing paper by OBA level may have validated this procedure.

Data Collection

An experiment was designed to collect data regarding paper brightness. Ten sheets each of eleven types of paper were collected for the research. The paper samples were protected by acetate sheet protectors during the measurement procedure, with the samples exposed only to take the readings. Care was taken to keep the paper from becoming dirty or scuffed during measurement.

After measuring the OBA level of each paper sample, the researcher measured brightness indices of each paper on every instrument twice a day (once in the morning, and again in the late afternoon) for five days. Data were collected from each paper sample in both the grain and cross-grain directions. The average of these two readings was calculated and used for the subsequent analysis. To minimize potential measurement variance, one researcher alone performed the readings, and the data from days three and four were utilized to assess consistency in the instruments and metrics and address the research questions. As it is recognized that variance can be introduced by the operator (e.g. Mandel, 1972), the approach allowed for an initial two-day period of readings to assure that the researcher was completely familiar with the respective instrument and the procedures. These days one and day two readings, together with readings from the fifth day, provided a type of insurance: in the event that anomalies were noted in day three and four readings the readings from the other days could be analyzed to make informed decisions regarding the data, while potentially avoiding the need to return to the measurement process.

Data Analysis

After collection, the data were analyzed. Boxplots and other graphics were used to graphically display possible variance. Because of the different operational characteristics of the instruments, some limitations of specific instruments were noted; these limitations influenced the subsequent analysis. For example, the X-Rite 530 is only capable of showing whole numbers under 100, and if the paper brightness reading is over 100 the instrument will display only "XXX" on the screen. Because of this limitation, the brightness index from the highest OBA sample utilized is not readable with X-Rite 530. As a result of that limitation, X-Rite 530 resulted in fewer readings than the other instruments.

Research Question 1

Before running the analysis for Research Question 1, data were visually analyzed for normality and outliers using a Q-Q test. Although results were not perfectly normal and some outliers were noted, it was determined that the outliers would remain in the analysis. To validate this decision, outliers were removed, and the tests were conducted again: in each case, the conclusions remained consistent.

After the normality test, a paired t test was performed and indicated that the difference in readings over two days did not elicit a statistically significant difference in TAPPI brightness, $t(109) = -1.79$, $p = 0.077$. The paired sample mean readings between the days analyzed were highly correlated ($p < 0.001$). For Research Question 1, recorded variances were therefore minimal, which underscores the consistency of this laboratory-grade instrument.

To visually examine the categories of OBA level, a boxplot of the variance in TAPPI brightness readings was generated and is reproduced in Figure 1. It is noted that the boxplot representing the 20 readings of the two papers with no OBA was similar to the 50 readings representing the five paper with low to moderate OBA in terms of mean and variance. These categories were combined to create a no to moderate OBA category for subsequent analysis. A boxplot displaying these two categories is reproduced in Figure 2.

Research Question 2

To determine how much variance can be expected when reading brightness indices, test-retest reliability was assessed through a paired samples t test; correlations were also reported. Day 3 and Day 4 brightness indices from each instrument were utilized for the analysis. The paired sample t test was performed and indicated that the difference in readings over two days did not elicit a statistically significant difference in brightness indices, except for the specific case of the X-Rite 939, $t(109) = -3.119$, $p = 0.002$, as illustrated in Table 3. The specific case of the X-Rite 939 is curious and resurfaces in the analysis of Research Question 3.

Table 3

Brightness Indices Paired t test Correlations Results

Instrument	t	df	Sig.(2-tailed)
X-Rite 939	-3.119	109	0.002
X-Rite 530	-0.587	99	0.558
X-Rite eXact	1.890	109	0.061
X-Rite Spectroeye	1.820	109	0.072
Minolta FD-7	-0.799	109	0.426

For the sub-question, it is noted that distributions of the brightness values for the OBA groups were not similar, as assessed by visual inspection. The non-parametric Mann-Whitney U Test was therefore used for the analysis. There was a statistically significant difference in Brightness values between the OBA groups, as illustrated in Table 4.

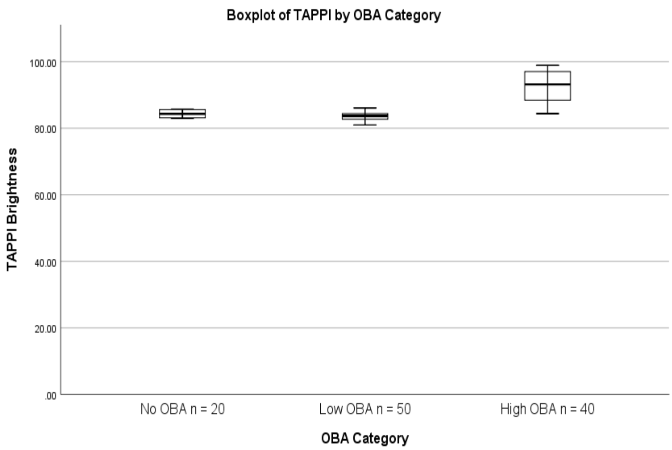

Figure 1: Boxplots of TAPPI Brightness by OBA Category: No OBA, Low to Moderate OBA, and High OBA

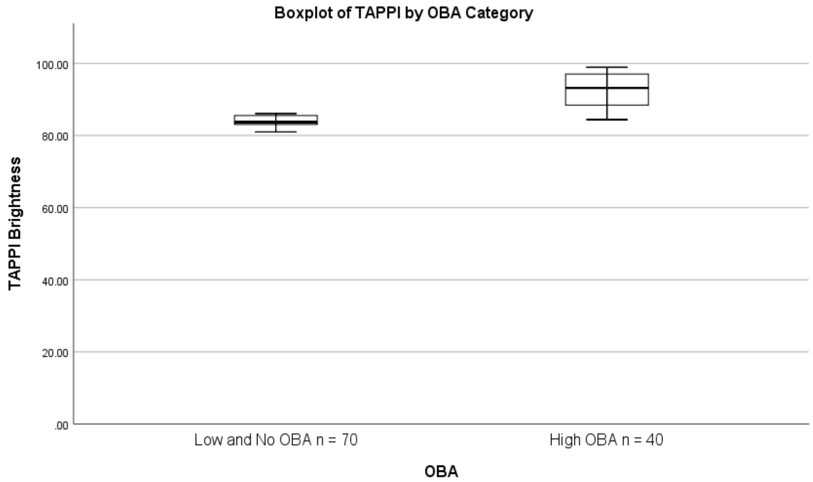

Figure 2: Boxplots of TAPPI Brightness by OBA Category.
No to Moderate OBA and High OBA.

Table 4
Brightness Indices Mann-Whitney U Test Result Table

Instrument	U	z	p	High OBA Mean Rank	Low OBA Mean Rank
X-Rite 939	300	-5.64	<0.001	75.50	39.79
X-Rite 530	243	-6.08	<0.001	77.40	38.97
X-Rite eXact	167	-6.64	<0.001	79.93	37.89
X-Rite Spectroeye	119	-7.00	<0.001	81.53	37.02
Minolta FD-7	679	-2.79	<0.001	62.87	45.02

Research Question 3

For research question 3, systematic bias and proportional bias are utilized for the analysis. Systematic bias is an assessment of agreement with a so-called "gold" standard; in this instance, the standard is the TAPPI brightness readings from the Technidyne laboratory-grade instrument. An independent samples t test is used to calculate the difference between TAPPI brightness and brightness indices to examine potential systematic bias.
Proportional bias is an analysis of agreement throughout the range of measurements. Tukey mean-difference plots, otherwise known as Bland-Altman plots, were utilized to examine possible proportional bias, and possible outliers were also identified.

Systematic bias
The presence of systematic bias is indicated if the mean value of the difference differs significantly from zero on the basis of the independent samples t test. Results are indicated in Table 5; it is noted that the mean differences recorded were all statistically significant and that they were all negative values.

Table 5
Independent t test Results

Instrument	t	Df	Sig(2-tailed)	Mean Difference
X-Rite 939	-11.81	109	<0.001	-5.38
X-Rite 530*	-22.81	99	<0.001	-2.79
X-Rite Spectroeye	-19.97	109	<0.001	-2.42
Minolta FD-7	-16.05	109	<0.001	-1.87
X-Rite eXact	-54.13	109	<0.001	-6.29

* Brightniss Indes on the high-OBA Neenah Paper was not readable with X-Rite 530

With the Tukey mean-difference plot, the mean difference represents the estimated bias, and the standard deviation is a measurement of fluctuations around the mean. Graphical results are illustrated in Figures 3 through 7.

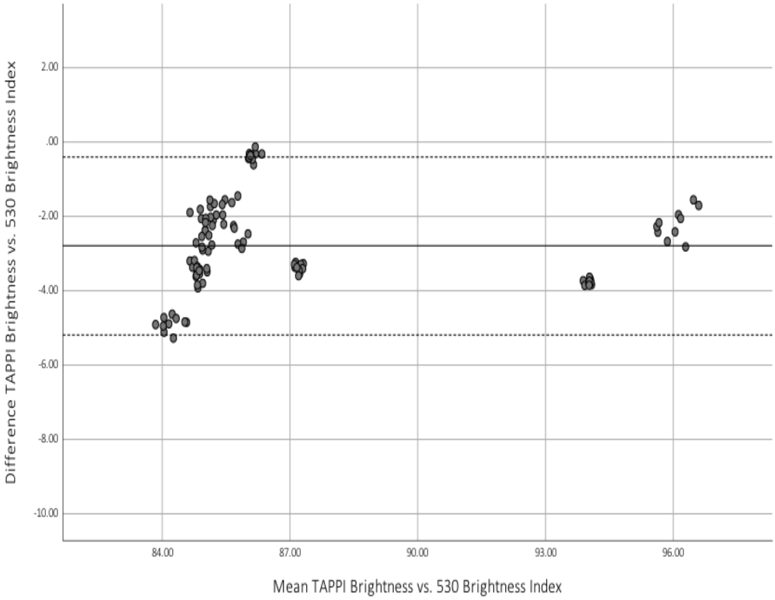

Figure 3. X-Rite 530 Tukey Mean-Difference Plot

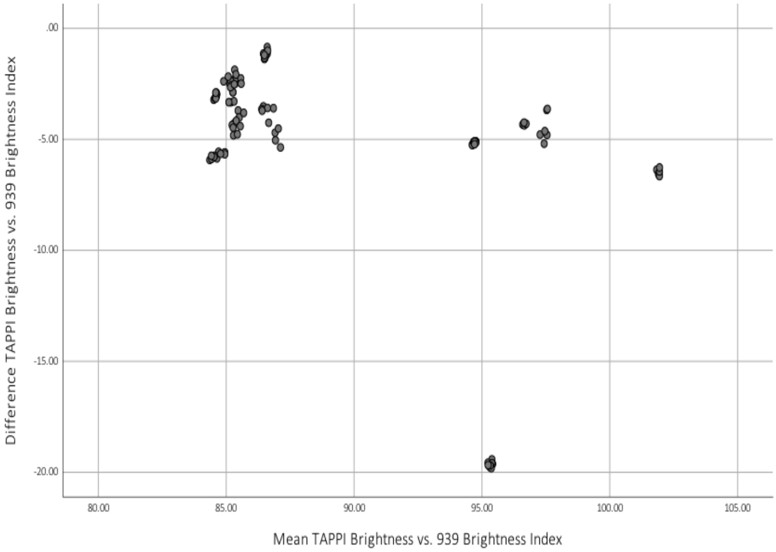

Figure 4. X-Rite 939 Tukey Mean-Difference Plot. Note cluster of readings at the bottom of the figure.

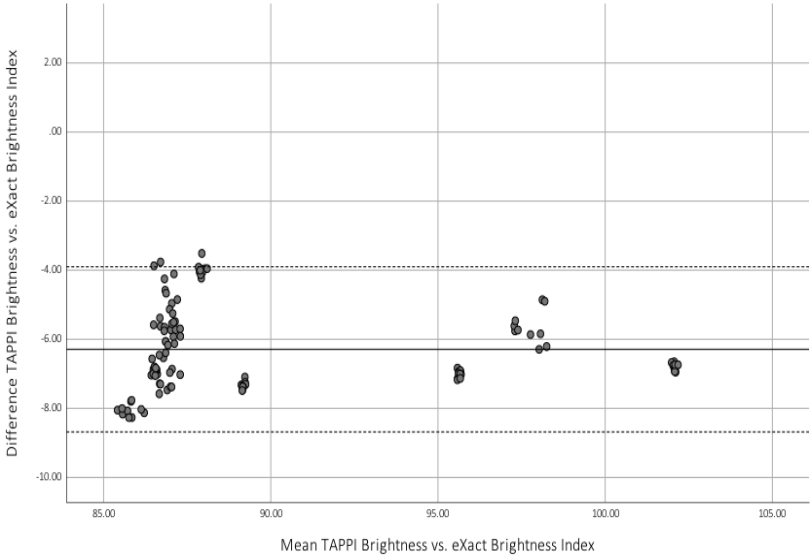

Figure 5. X-Rite eXact Tukey Mean-Difference Plot

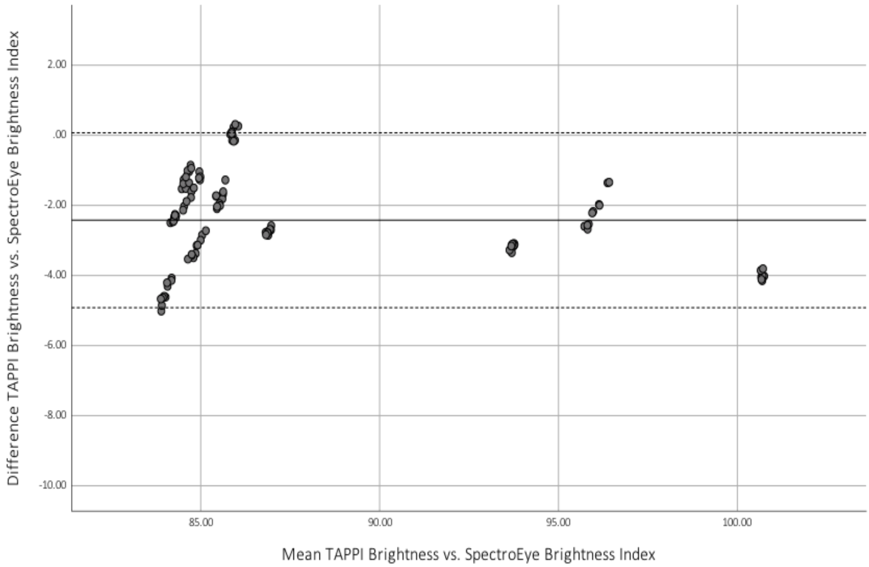

Figure 6. X-Rite SpectroEye Tukey Mean-Difference Plot

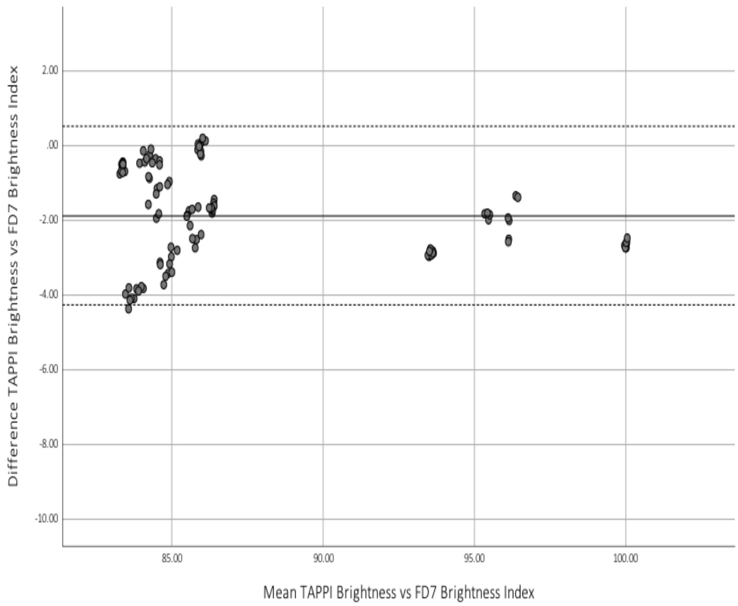

Figure 7. Konica Minolta FD-7 Tukey Mean-Difference Plot

A particular case in noted again with the X-Rite 939. Here, one particular substrate, SBS, represented an outlier. As illustrated in Figure 4, there was a notable cluster of brightness readings showing a difference of nearly -20 from TAPPI brightness. While this was curious, it was recognized that this particular model represents an older instrumentation design, especially when compared to the Konica Minolta FD-7 and the X-Rite eXact. The optical characteristics of the X-Rite 939 may have resulted in excessive variance due to the surface characteristics or other property of the SBS substrate; therefore, these particular readings were removed from the analysis. The resultant Tukey mean-difference plot is shown in Figure 8. Table 5 displays an updated t test for systematic bias with the problematic substrate removed from the X-Rite 939 readings: here, the results obtained by the X-Rite 939 were more consistent with the other instrumentation.

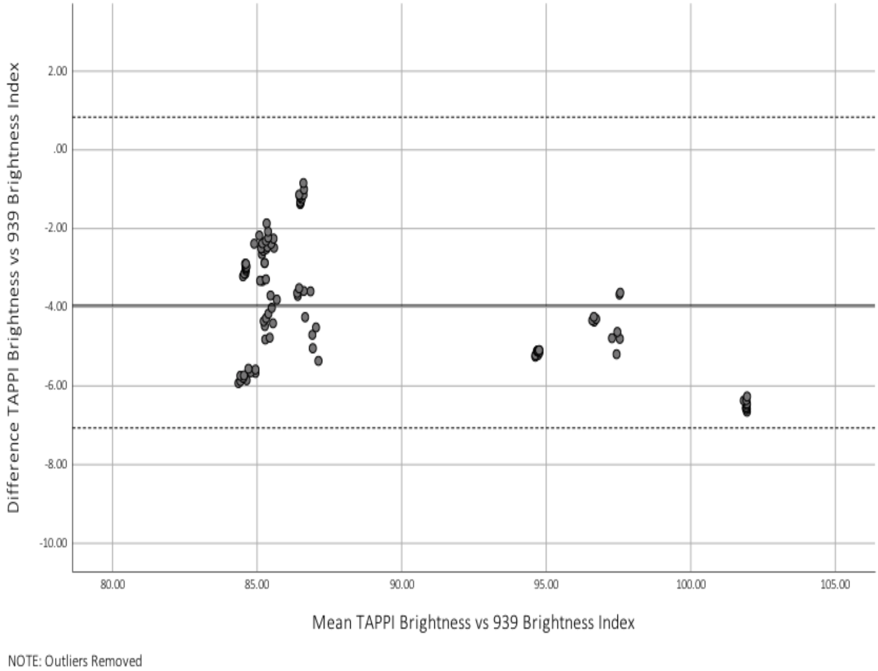

Figure 8. X-Rite 939 Tukey Mean-Difference Plot with Outliers Removed

Table 5
Independent t test Results, SBS removed from X-Rite 939 readings

Instrument	t	Df	Sig(2-tailed)	Mean Difference
X-Rite 939*	-24.88	99	<0.001	-3.95
X-Rite 530*	-22.81	99	<0.001	-2.79
X-Rite Spectroeye	-19.97	109	<0.001	-2.42
Minolta FD-7	-16.05	109	<0.001	-1.87
X-Rite eXact	-54.13	109	<0.001	-6.29

*Problematic SBS readings removed from X-Rite 939, Brightness Index on the high OBA Neenah Paper is not readable with X-Rite 530

Discussion and Conclusions

The present study examined expected variance in reading TAPPI brightness and brightness indices across a range of printing papers. Other than brightness, the presence of OBAs was the only other characteristic examined; these were operationalized through the OBA-Check feature of the Techkon SpectroDens. While this methodology does introduce limitations, practical and relevant conclusions can be observed from this work.

The first two research questions examined the variance that can be expected when using the selected instrumentation to measure brightness. In the case of the laboratory-grade Technidyne meter, the paired sample t test indicated that the difference in readings over a two-day period did not elicit a statistically significant difference in TAPPI brightness as the paired sample mean readings between Day 3 and Day 4 were highly correlated. Similarly, for Research Question 2, a paired sample t test resulted in the observation that the difference in readings over two days did not elicit a statistically significant difference in brightness indices, except for the specific case of the X-Rite 939. As previously discussed, the 939 was problematic with one of the substrates when compared with the other utilized instruments.

To ascertain the effect of OBAs with the laboratory-grade Technidyne instrument, as examined in sub research question 1, and the production-grade instruments, as examined in sub research question 2, the non-parametric Mann-Whitney U test indicated that variances among the brightness indices for the instruments as noted. Users are therefore cautioned about the observed increase in variance with higher OBA papers when measuring brightness.

These findings suggest that users of any of these instruments, save the X-Rite 939, can confidently realize repeatable readings, particularly with low- and moderate-level OBA substrates. However, future researchers may want to conduct full Gage R&R studies, which could serve to validate further the results obtained here.

In addition to the concerns mentioned above with the X-Rite 939 in this context, it is also important to recognize the limitations of the X-Rite 530 versus the other instruments. As indicated, the model X-Rite 530 does not display brightness readings over 100, so it was impossible to record readings for the substrate selected that exhibited the highest level of OBAs. Further, the X-Rite 530 only displays whole numbers. First introduced in 1997, the model X-Rite 530 represents the oldest technology of all of the instruments analyzed, and when sold as new, was offered at the lowest cost. The limitations of older instrumentation noted here should serve to caution practitioners in this domain.

It is relevant to note that both the X-Rite 530 and X-Rite 939, although popular in the field, were introduced before ISO 13655:2009. This ISO standard defines measurement conditions for color measurement instrument manufacturers with the goal of achieving better agreement between visual assessment and measurements (ISO, 2009). While it is also true that the X-Rite SpectroEye used in the present study was introduced prior to 2009, the X-Rite eXact and the Konica Minolta FD-7 utilized here comply with ISO 13655:2009. It is recognized that while some of the older instruments may be limited, from a practical standpoint, moving ahead such legacy instrumentation will likely be utilized less and less, especially as factory service and certifications are retired.

Turning to the third, and primary research question, the data suggest that of all of the production-grade instruments evaluated, the Konica Minolta FD-7 was closest to the laboratory-grade TAPPI meter, and that the X-Rite eXact was the furthest. All of the production-grade instruments read higher brightness values than the TAPPI standard, and all were significantly different than the TAPPI readings obtained with the laboratory-grade Technidyne brightness meter.

In examining the difference between high- and no to moderate OBA papers visually using the Tukey mean difference plots, the variances were generally similar across all production-grade instruments analyzed. As discussed, the X-Rite 939 represented an exception here. Some of the no to moderate OBA substrates seemed to increase the standard deviations of the brightness index readings. This suggests a limitation in the OBA categorization methodology that was used here: future researchers may choose a different methodology to separate papers by OBA presence. Furthermore, as OBA was the only paper characteristic analyzed by the present study, future researchers may choose to evaluate other factors, such as surface characteristics (e.g., smoothness, gloss) to better determine how these indices correlate to the standard.

Paper manufacturers are encouraged to indicate the specific brightness metric utilized and the brightness measurement technique (i.e.: grain direction, cross grain direction, average of the two) with paper label information. Meanwhile, consumers are cautioned to be advised of advertised brightness differences depending on the metric utilized.

Reference List

1. Cepova, L., Kovacikova, A., Cep, R., Klaput, P., & Mizera, O. (2018). Measurement system analyses - gauge repeatability and reproducibility methods. Measurement Science Review, 18(1), 20-27. doi: http://dx.doi.org.ez-proxy.rit.edu/10.1515/msr-2018-0004

2. IPSTESTING. (2017). TAPPI T452 Brightness. 1. Retrieved from https://ipstesting.com/find-a-test/tappi-test-methods/tappi-t-452-brightness/

3. ISO. (2009). ISO 13655:2009. Retrieved from https://www.iso.org/standard/39877.html

4. Mandel, J. (1972). Repeatability and Reproducibility. Journal of Quality Technology, 4(2), 74-85. doi:10.1080/00224065.1972.11980520

5. Sloop, R. (2009). Understand gage R&R. Quality, 48(9), 44-47. Retrieved from http://search.proquest.com.ez-proxy.rit.edu/docview/235257226?accountid=108

6. Techkon USA (2018). SpectroDens Manual. Retrieved from http://www.techkonusa.com/wp-content/uploads/2017/04/SpectroDens-Manual-Web.pdf

Bruce Myers

Hanyi Cheng

George Shields, M.Sc., UPM North America, Director of Technical Customer Service

george.shields@upm.com

Alexandra Pekarovicova, PhD., Professor, Western Michigan University, Kalamazoo, MI

a.pekarovicova@wmich.edu