Clin Oral Invest (2005) 9: 36–42
DOI 10.1007/s00784-005-0014-3

ORIGINAL ARTICLE

Burkard Hugo · Tobias Witzel · Bernd Klaiber

Comparison of in vivo visual and computer-aided tooth shade determination

Received: 20 September 2004 / Accepted: 22 August 2005 / Published online: 8 October 2005

Abstract The aim of this study was to evaluate the performance of shade-determining devices. For the daily practitioner, it is essential to know whether modern computer-aided shade selection is reliable in everyday life. So the question of how the clinical usability of these machines could be rated has to be clarified. In the following, three actual devices available in the market were compared using a human observer's perception. The SpectroShade device (MHT Optic Research AG, 8155 Niederhasli, Switzerland), the ShadeVision device (X-Rite Co., Grandville, USA) and the Digital Shade Guide DSG4 (A. Rieth, 73614 Schorndorf, Germany) were assessed with respect to their agreement with the color perception of three examiners looking at 57 test persons (six teeth each for a total of 342). Shades were reported in Vita Classical shades. It could be demonstrated that every single human examiner showed a significantly higher agreement value (human group on average 40.2%) when compared with the remaining five methods than each computer-aided tooth shade determination device. The devices reached on average only a value of 28.6%, whereas the X-Rite ShadeVision showed a significant better result (33.2%) than the MHT SpectroShade and Rieth DSG4 (27.0 and 25.7%). Identical shade results given by all three methods of a group (group of three devices and three humans) were found to be rather low for the computer-aided devices (9.9%) compared with humans (36.7%). All six methods together agreed in 3.3% of the cases. It becomes evident that the methods—especially the computer-aided shade determination—are rather divided about the respective tooth color. Deficiencies of the instrumental as well as the visual detection become obvious. The best agreement level was performed by the human examiners. The best agreement of the evaluated devices was obtained—generally as well as among the human testers—by the X-Rite ShadeVision system, followed at a statistically significant distance by the MHT SpectroShade and the Rieth DSG4. The agreement among the examiner group was 52.9%, significantly better than that of each device compared to this group (31.3% on average). Color detection and its realization are very complex. As shown, in many cases, computer-aided color shade determination of natural teeth seems to not reflect human perception.

Keywords Tooth color · Shade selection · Measurement · Agreement · Color perception

Clinical Relevance
For color determination of natural teeth, there is significant disagreement between the shade selected by human perception and the shade obtained with computer-aided instrumentation.

B. Hugo (✉)
Department of Orthodontics, Bavarian Julius Maximilian University of Würzburg,
Würzburg, Germany
e-mail: burkard.hugo@mail.uni-wuerzburg.de
Fax: +0049-0931-20173300

B. Hugo · B. Klaiber
Klinik und Polikliniken für Zahn-, Mund- und Kieferkrankheiten, Universität Würzburg,
Abteilung Kieferorthopädie, Pleicherwall 2,
97070 Würzburg, Germany

T. Witzel
Gartenstr. 1,
97072 Würzburg, Germany
e-mail: mail@tobias-witzel.de

B. Klaiber
Department of Operative Dentistry and Periodontology,
Bavarian Julius Maximilian University of Würzburg,
Würzburg, Germany

Introduction

Even today, the visual shade selection by the dentist, the dental assistant or the dental technician is still the most common method to adapt a tooth restoration to the natural tooth color. Next to form and function, the harmony between natural and artificial teeth, fillings or facings is a major sign of quality. The development and potentialities of modern ceramic and composite restoration materials have given rise to increased aesthetic demands of the patients that often cannot be met due to a lack of objectivity in shade selection. In 1970, Culpepper [5] already demonstrated visual shade selection is often unreliable and imprecise. The impression of color by the bare eye is rather subjective but not objective or even measurable and is therefore perceived and felt differently by every individual [5, 13]. Billmeyer and Saltzman [2] define the term 'color' as 'the result of the physical modification of light by colored objects that is perceived by the human eye and interpreted by the brain'. Seeing therefore consists of rather different processes, including physical ones like light and reflection, physiological ones in the eye and psychological ones in the brain of the observer. Each of these implements variable factors such as environment, illumination, angle of perception, restricted sample space, tiredness of the eye, color perception deficiencies a.s.f., not permitting a uniform and objective color vision [11]. These uncertainties of visual color determination therefore make an objective method desirable. By using instrument-based color measurement (colorimetry), it is generally possible to exclude such error sources and thus obtain neutral reproducible results. However, when looking at the layered structure of a tooth consisting of enamel, dentine, cement and pulp, it becomes obvious that within the solid substance of a tooth, a number of bordering surfaces determine the complex visual impression as well as its surrounding soft tissues. Therefore, dentistry poses special requirements to the instrumental technology different from that in industries where color measurement has been a well-established method for many years. Because of this, tooth color measurement is as yet unable to find a definite place in the routine application in dental practice. Since the end of the 1990s, an increasing number of computer-based instruments for shade selection have entered the market (Table 1). The manufacturers promise to the user a clear-cut shade selection and thus more contented patients. The subject of this study was to examine the ability of such devices to agree with the human color sensation, making the high investment in a computer-based system attractive.

Materials and methods

In this study, color distribution and agreement of three tooth color shade detectors available on the market were determined:

- SpectroShade (MHT Optic Research AG, 8155 Niederhasli, Switzerland, http://www.mht.ch, Software V. 2.20D)
- ShadeVision (X-Rite Co., Grandville, USA, http://www.shadevision.com, Software V. 1.20)
- Digital Shade Guide DSG4 (A. Rieth, 73614 Schorndorf, Germany, http://www.a-rieth.de, Software V. 1.7)

The devices were compared with three human examiners looking at 57 test persons (six teeth each for a total of 342) under clinical condition.

Test persons

The subjects of this investigation were the maxillary central and lateral incisors and canines of 57 test persons (30 women and 27 men, average age 31±9 years). The maxillary front was chosen for its visual exposition and therefore highest requirements in aesthetics. Over half of the persons looked at (54%) were in the age group between 24 and 27 years. Potential test persons having restorations on the teeth were excluded from this study. The test persons were all dental students or dental personnel of the ZMK-Klinik, University of Würzburg. Due to the good oral hygiene that was found without exception, there were no exogenous solid deposits or stain visible. All the test persons had received a professional tooth cleaning not longer than 6 months ago.

Visual shade selection by human observers

Before using the instruments, the tooth shades were determined visually by three examiners who had been

Table 1 Devices available in the market

Unit	Manufacturer	Website	Launch
Shade-Eye-NCC	Shofu Dental Inc., San Marcos, USA	http://www.shofu.com	1997
Dental Color Analyzer	Wolf Industries Inc., Vancouver, USA	http://www.clearlight.com/~aei	1998
PR-650 SpectraScan	Photo Research Inc., Chatsworth, USA	http://www.photoresearch.com	1999
ShadeScan	Cynovad, Montréal, Canada	http://www.cynovad.com	2000
ClearMatch (software only)	Clarity Dental Corp., Salt Lake City, USA	http://www.claritydental.com	2001
Digital Shade Guide DSG4 (+)[a]	A. Rieth, Schorndorf, Germany	http://www.a-rieth.de	2001
Ikam	Metalor Technologies Dental Division, North Attleboro, USA	http://www.metalor-dental.com	2001
ShadeVision/ShadeRite Dental Vision[a]	X-Rite Inc., Grandville, USA	http://www.shadevision.com	2001
SpectroShade[a]	MHT Optic Research AG, Niederhasli, Switzerland	http://www.mht.ch	2001
Vita Easyshade	Vident, Brea, USA	http://www.vident.com	2002

[a]Devices examined in this study

tested for color perception deficiencies (Farnsworth test, panel D15 desaturated).

The three testers determined the color shades independently and with no time limits for each tooth third (cervical, middle, incisal) and also the whole tooth (subjective average) using the same Vita Classical shade guide (Vita Zahnfabrik, H. Rauter Gmbh & Co. KG, Bad Säckingen, Germany) that was regularly ordered into the four common shade groups—A, B, C and D. The observers consisted of two experienced dental technicians and one dental student in his last training semester. The examiners were between 26 and 38 years old. The lighting of the test area came from four daylight fluorescent light tubes of 36 W/5,000 K each, conforming to the standard ISO 3664, made especially for color comparison purposes (Just Normlicht, 73235 Weilheim/Teck, Germany). The light sources were mounted in two reflector cases (Kaiser, model RB 5004, 74711 Buchen, Germany). The illumination of the test area was done at a 45°-angle from frontally right and left. The distance between light source and object was adjusted such that 1500 lx could be measured. Weak daylight incident from the front was superimposed by the described light tubes. The persons to be tested were requested to remove their make-up and to brush their teeth for about 1 min (one-way toothbrush, Hager & Werken GmbH & Co KG, 47269 Duisburg, Germany) in order to eliminate soft deposits. Shades were selected with slightly moistened teeth. Colorful dress was covered with a grey cloth to exclude possible interference. The Vita Classical shade guide was used as the basis for this study because of its wide distribution and the testers' familiarity with it.

Instrumental shade determination

The computer-aided shade selections were always done by the same person. Prior to this study, the examiner carried out substantial training in order to become familiar with the handling of the particular equipment.

At the beginning of each session, the instruments were allowed to warm up for 15 min and were calibrated according to manufacturer's specifications using the included standards. The test persons were instructed to lean their heads against the headrest of the treatment chair during the measurement and to keep their mouth slightly opened. The tongue had to be in a relaxed position since pressing of the tongue against the maxillary front might cause mismeasurements by tissue shining through due to incisal tooth translucency. The surgical lamp of the treatment unit remained turned off during measurements. All measurements were checked generally by a second procedure to ensure measurement errors.

MHT SpectroShade system

In using the MHT SpectroShade instrument (Fig. 1), the measuring head, which is connected to the computer by a light-conducting data cable, was positioned with its mouthpiece on the alignment. The central position of the tooth to be measured was checked and corrected on the computer screen. As required by the manufacturer, the examiner carried out the measurements twice in a row on the slightly moistened teeth and subsequently compared the color distance to exclude mismeasurements (default value by the manufacturer's manual: $\Delta E<1$). Particular attention—as with all instruments—was given to the lips, which when squeezed in could render the results invalid. After the measurement was triggered at the measuring head, the tooth area under scrutiny was recognized and marked automatically by the software. The corresponding data were stored automatically.

X-Rite ShadeVision system

After the X-Rite ShadeVision handset (Fig. 2) was taken out from the docking station, the required tooth (also slightly moistened) was selected from the tooth scheme shown on the black-and-white touch screen liquid crystal display. The instrument was positioned such that the highlights showed in the medial cervical third of the tooth (referring to manufacturer's instructions/manual). Finally, the measurement was triggered on the screen, and the next tooth was looked at. At the end of the measurements, the data were transmitted to the computer. The tooth area from which the shade was to be taken was determined manually using the software (wizard tool), and the output was saved.

Fig. 1 MHT SpectroShade unit with Spectrophotometer integrated into the computer casing

Fig. 2 X-Rite ShadeVision handset (colorimeter) with universal serial bus (USB) docking station

Rieth Digital Shade Guide DSG4

The pin-shaped Rieth DSG4 measuring head was placed on the surface of the teeth as perpendicularly as possible, beginning with the right maxillary canine. The head was connected to the colorimeter by a light-conducting cable. The Rieth DSG4 hardware (Fig. 3) was connected to the PC by standard RS-232 serial interface. The shades were determined for each tooth in its cervical, middle and incisal third. As requested by the manufacturer, the measuring tip and the tooth were moistened with water at every measurement in order to suppress the influences of surface structures (e.g. dullness, lustre).

The measurement was triggered with the foot switch. Values were stored by the software in the tooth scheme and the respective third of each tooth. Compared with the MHT SpectroShade or X-Rite ShadeVision devices, the examiner had to make three separate measurements to obtain the desired output.

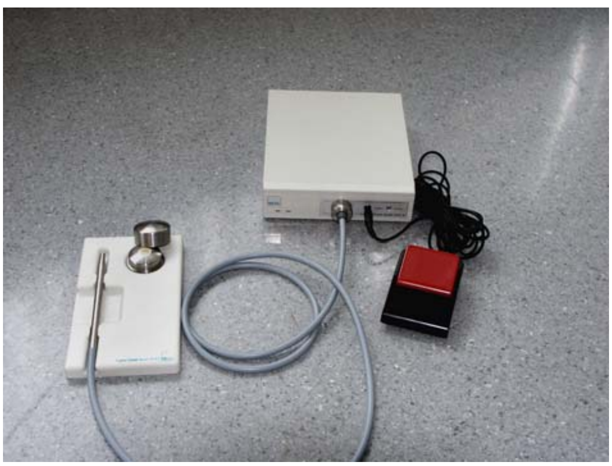

Fig. 3 Rieth DSG4 unit: colorimeter, foot switch, pin-shaped measuring tip with rack and calibration pattern

Evaluation

As already mentioned, the colors of the tooth were determined using the Vita Classical shade guide, which is widely used in dentistry and therefore the de facto standard. In the following examination, the average tooth color was used if possible. For the Rieth DSG4 device, data output of the central third had to be used for analysis for technical reasons (no possibility of determining the color of a larger area/whole tooth). Concerning the X-Rite ShadeVision software, the given output for the whole tooth was identical with the values given as ‘middle third’; the MHT SpectroShade system was the only one that computed the ‘all over color’ from all given values.

Furthermore, the MHT SpectroShade system offers the special option of three different color formulas to calculate color distance. In our case, the preset formula ‘dE’ (‘delta E’) was used because it represents the generally accepted industrial standard [1]. The other optional formulas, ‘dE-MHT’ and ‘dL’, would give a stronger weight-to-brightness difference in shade determination.

The X-Rite ShadeVision instrument signalled in some cases that the chosen color was not contained in the Vita Classical color space investigated and would therefore be replaced by the nearest fitting shade. Since the other instruments had to make a decision as well between the 16 possible Vita shades, these messages were ignored by the investigator. Sometimes the X-Rite ShadeVision suggested intermediate shades (such as A1-A2). In such cases, the first mentioned shade was used for analysis.

The Rieth DSG4 instrument furthermore quantified intermediate shades, e.g. 80% A1 and 20% A2, in which case the larger part was taken as the result. In a few cases of exactly 50% each, the first given shade was used for the evaluation.

In this analysis, only qualitative statements can be made about the agreement of the methods because L^*a^*b values —optimal for colorimetry—are only outputted by the MHT SpectroShade device. Furthermore, according to its manufacturer, the L^*a^*b values would only be suitable for comparison within the system. Therefore, comparisons with other L^*a^*b sources would not be admissible.

Results

Agreement amongst the methods and with the ‘majority opinion’

Figure 4 gives an overview of the overall general agreement (same shade selected) of one method with all the five other methods (summation of separate pair-agreements). The X-Rite ShadeVision at 33.2% scores significantly better in agreement (chi squared test according to Pearson, level of significance $p \leq 0.5$; always used in following statistical comparisons if not mentioned separately) than the MHT SpectroShade and Rieth DSG4 (27.0 and 25.7%, respectively). Accordingly, the group of instruments correlated on average in 28.6% (i.e. average of the devices in

Fig. 4), and the group of human examiners showed an average correlation of 40.2% (i.e. average of the examiners in Fig. 4). The two groups deviate from each other significantly.

The same shade was obtained by all three instruments in 9.9% and by all human examiners in 36.7%. The agreement rate of all six methods was 3.3% (Fig. 5).

None of the in vivo tested teeth was a sample tooth with known vita color. Thus, there is no gold standard, and the true color of the in vivo teeth could not be determined with certainty. Therefore, a 'majority opinion' was defined as the color of every tooth that was determined by the majority (the most frequently given shade) of the six methods. Teeth for which no clear-cut majority could be obtained (e.g. if six several shades or the same amount of unequal shades were given) were not taken into consideration. This was the case for 14.8%. No evaluations using statistical methods containing the 'majority opinion' could be carried out because the 'majority opinion' is composed of the data with which it is to be compared.

Figure 6 therefore shows a concluding (non-statistical) comparison of the individual methods with the 'majority opinion'. Amongst the devices, here the X-Rite ShadeVision shows the highest rate of agreement with 55.9%, MHT SpectroShade came second best with 43.9%, followed by the Rieth DSG4 with 41.1%. The examiners dental student, dental technician 1 and dental technician 2 have agreements with the majority opinion of 70.0, 70.9 and 67.7%.

Among the three possible pairwise combinations of the instruments, the agreement of MHT SpectroShade vs. X-Rite ShadeVision was 28.4%, X-Rite ShadeVision vs. Rieth DSG4 25.9% and MHT SpectroShade vs. Rieth DSG4 19.1%. An agreement overview of all possible method combinations is shown in Fig. 7.

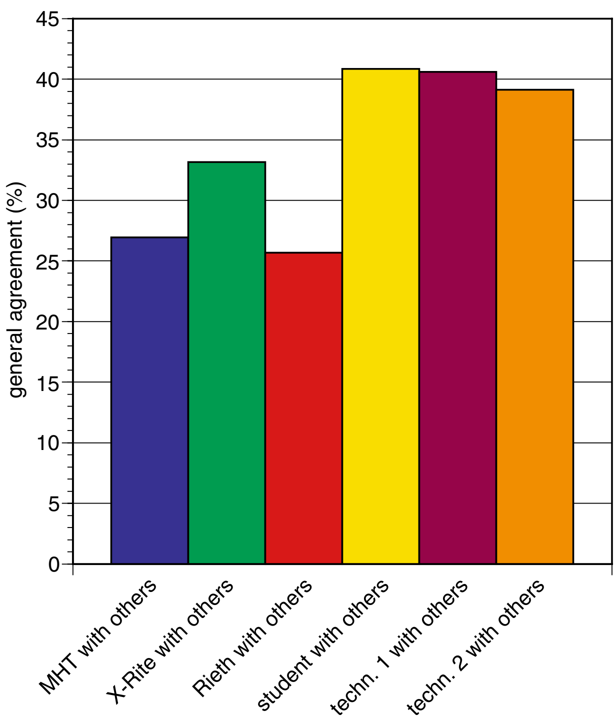

Fig. 4 Agreement of the individual methods with the others (independent of tooth third and type)

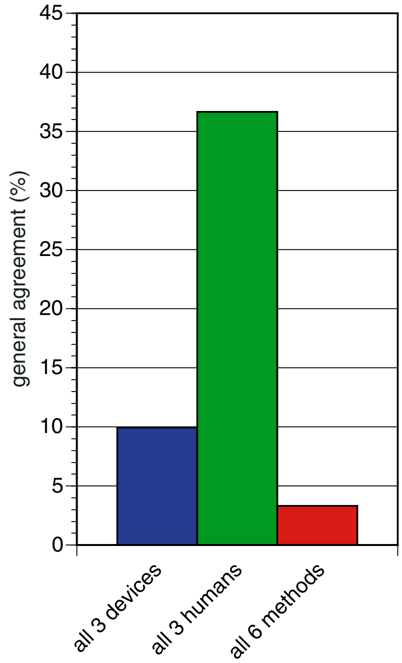

Fig. 5 General agreement of all methods in the group

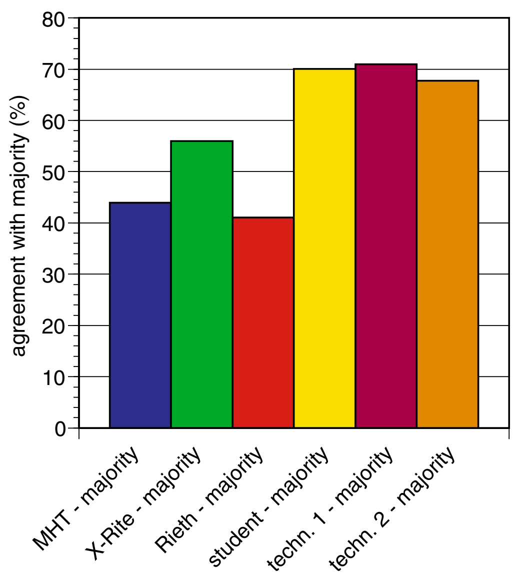

Fig. 6 General agreement of the methods with the 'majority opinion'

Agreement of the devices with the group of human examiners

The agreement rate of the individual instrument with the group of human examiners (summation of the agreement of the three pairs, i.e. instrument–human 1, instrument–human 2 and instrument-human 3) was 29.1% for the

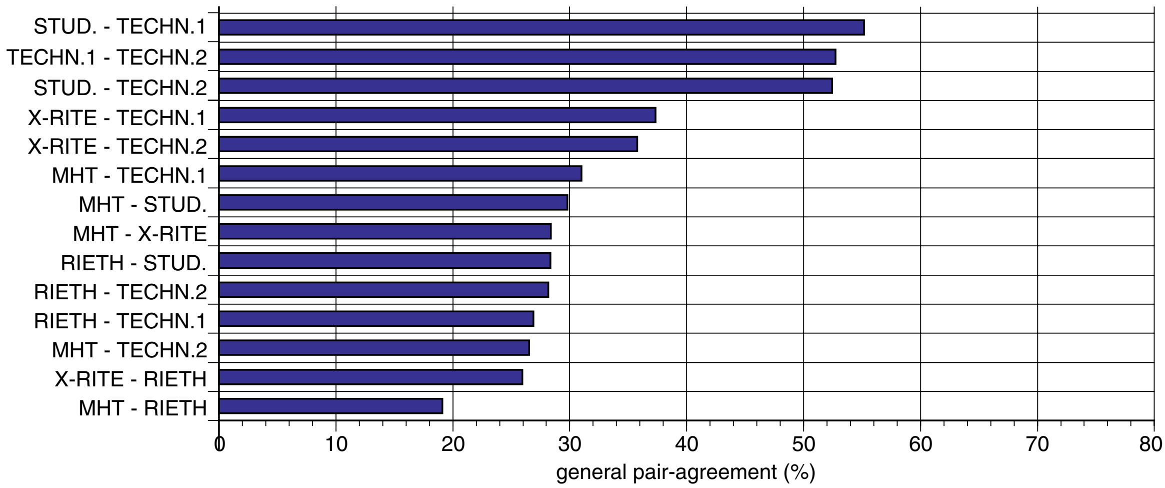

Fig. 7 Agreement rates of all method combinations

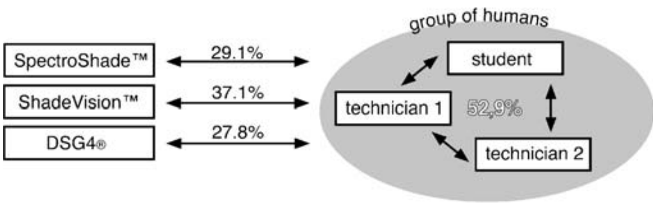

Fig. 8 Agreements with the human group, overview-diagram

MHT SpectroShade unit, 37.1% for the X-Rite ShadeVision and 27.8% for the Rieth DSG4. The pairwise agreement between the three human examiners averaged at 52.9% (Fig. 8).

The agreement rate of each individual device with the examiner group was shown to be significantly lower than the examiners among each other.

In comparison, the agreement rates of the devices with the human group shown above also differ significantly among each other (chi squared test according to Pearson, alpha adjusted to p=5/3=1.67 according to Bonferroni); therefore, the X-Rite ShadeVision unit shows a significantly higher agreement rate than MHT SpectroShade and Rieth DSG4. MHT SpectroShade and Rieth DSG4, however, do not differ significantly (Fig. 8).

Discussion

Agreement of the methods amongst each other and with the 'majority opinion'

Looking at the general agreement of the methods with the others, human examiners show a significantly higher value than the instruments. In turn, the X-Rite ShadeVision shows a significantly better agreement with the other methods than the MHT SpectroShade and Rieth DSG4 units.

The cases in which the group of three instruments was of the same opinion (result MHT SpectroShade=result X-Rite ShadeVision=result Rieth DSG4) pertain to values that lagged behind the author's expectations (Fig. 5). While the human examiners agree unanimously upon every third tooth (36.7%), this is only true for every tenth (9.9%) among the instruments! This result impressively underlines the inadequacies of shade determination, whether carried out by humans or even instruments.

It is interesting to note in the evaluation of the agreements that all methods always correlate best with the 'majority opinion'. This means that there is no 'method' representing the color of a tooth better than the 'majority opinion', although as pointed out before, this cannot be taken as a gold standard.

Among the three possible pair combinations of instruments, the agreement of MHT SpectroShade with X-Rite ShadeVision is better than that of X-Rite ShadeVision with Rieth DSG4 and is again better than that of MHT SpectroShade with Rieth DSG4.

This ranking is continued in the comparison of the individual methods with the 'majority opinion' (Fig. 6); human examiners show better results than the instruments.

Amongst the instruments, the X-Rite ShadeVision comes out top, followed by MHT SpectroShade and Rieth DSG4. The differences in agreement of the MHT SpectroShade with Rieth DSG4 units are clearly smaller than that with the X-Rite ShadeVision.

Furthermore, the results of the human group are within the frequencies of 52 to 80% described by Faber and Schlegel [6]. In their experiment, 188 persons of differing color experience were required to correlate identical Vita color samples with each other. Horn et al. [7] describe for their observers an agreement of 50 to 65% with the 'majority opinion' using only six Vita Classical shade tabs in selection on extracted teeth.

Agreement of the instruments with the human group

When the rate of agreement between each computer-aided device and the human group is statistically compared to the average rate of agreement of the three human examiners with each other, the following statements can be made.

The agreement rate of the three devices with the examiners is significantly lower than that of the humans amongst each other. The ability of the humans to agree with each other is thus proven to be greater than between the instruments and examiners.

The instruments also offer significant statistical differences with respect to their rate of agreement. The X-Rite ShadeVision unit shows a significantly higher rate with the human group than the MHT SpectroShade and Rieth DSG4 instruments, the two latter ones not being significantly different in this respect.

Wee et al. [12] evaluated in their in vitro experiments a 46.7% agreement in visual comparison between extracted teeth and color samples, manufactured by using another color detection device (ShadeEye-EX, Shofu Dental Inc., San Marcos, USA). In contrast with this analysis, theirs contained the laboratory production process as well, but was also obtained under laboratory conditions and using a different approach. Therefore, it can be constituted that they find a higher value than this study, demonstrating values between 27.8 and 37.1%.

The examined instruments show that—in principle—it is possible to develop systems equal or even superior to human vision when measured in the laboratory. However, there is a large deficit in the conversion of colorimetric data into color shades adequate to human perception. Faber and Schlegel [6], studying the colorimeter Castor, also had to state this, as well as Yap et al. [14] using a Minolta colorimeter.

Conclusion

This study was done to evaluate the performance of shade-determining devices in vivo because for the daily practitioner, it is more useful to know whether the computer-aided shade selection is reliable on his treatment chair. Recent in vitro studies [8, 10] showed good results for the MHT SpectroShade, Rieth DSG4 and X-Rite ShadeVision devices in reproducing values. Chu and Tarnow [3] as well as Chu [4] in exemplary cases also reported satisfactory results in producing crowns by using the MHT SpectroShade and the Shofu ShadeEye-EX (not included in this research) for color determination.

These positive results cannot be transferred to practical use. As shown above, there seems to be a significant disagreement of the devices with human color perception. Therefore, in the eyes of the authors, mismatching has to be expected in clinical application using shade-determining devices without following visual control.

With the title of his publication ‘Shades of a color. Illusion or reality?’, Mayekar [9] highlights the complex problem of the treated subject. In the end, however, the critical question has to be asked whether even an optimized software could clearly detect the complex layers, structures and translucencies of a tooth by colorimetric means or whether measuring technology has not met its physical limits.

Color detection and its realization are a very complex and wide field, which prevents proposing all aspects of shade determination in this essay. Therefore, in the near future, the authors want to complete this paper by further investigating the reproducibility of visual and computerized shade selection under standardized conditions.

References

1. Berger-Schunn A (1994) Praktische Farbmessung: ein Buch für Anfänger, eine Gedächtnisstütze für Könner, 2nd edn. Muster-Schmidt, Göttingen Zürich
2. Billmeyer FW, Saltzman M (1981) Principles of color technology, 2nd edn. Wiley, New York
3. Chu SJ, Tarnow DP (2001) Digital shade analysis and verification: a case report and discussion. Pract Proced Aesthet Dent 13:129–136 (quiz 138)
4. Chu SJ (2002) Precision shade technology: contemporary strategies in shade selection. Pract Proced Aesthet Dent 14:79–83 (quiz 84)
5. Culpepper WD (1970) A comparative study of shade-matching procedures. J Prosthet Dent 24:166–173
6. Faber FJ, Schlegel G (1999) Einfluß von Farbtüchtigkeit, Geschlecht und Beruf auf die Zahnfarbauswahl. Dtsch Zahnärztl Z 54:281–283
7. Horn DJ, Bulan-Brady J, Hicks ML (1998) Sphere spectrophotometer versus human evaluation of tooth shade. J Endod 24:786–790
8. Kohlmeyer B, Scheller H (2002) Zahnfarbenbestimmung mit dem PC-gestützten Messgerät “Digital Shade Guide”. Dtsch Zahnärztl Z 57:172–175
9. Mayekar SM (2001) Shades of a color. Illusion or reality? Dent Clin North Am 45:155–172
10. Paul S, Peter A, Pietrobon N, Hammerle CH (2002) Visual and spectrophotometric shade analysis of human teeth. J Dent Res 81:578–582
11. Sproull RC (2001) Color matching in dentistry. Part I. The three-dimensional nature of color. J Prosthet Dent 86:453–457
12. Wee AG, Kang EY, Johnston WM, Seghi RR (2000) Evaluating porcelain color match of different porcelain shade-matching systems. J Esthet Dent 12:271–280
13. Wyszeki G, Stiles WS (1982) Color science, 2nd edn. Wiley, New York
14. Yap AU, Sim CP, Loh WL, Teo JH (1999) Human-eye versus computerized color matching. Oper Dent 24:358–363