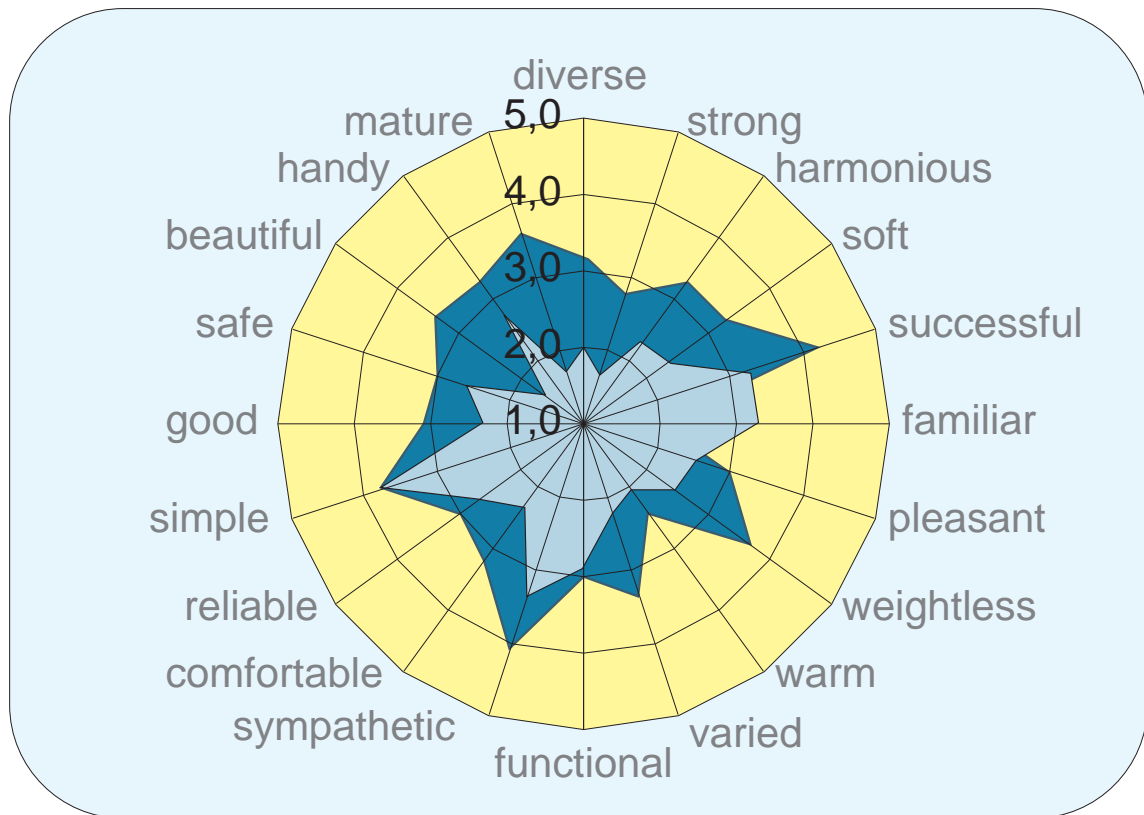


TESTVERFAHREN für Displayqualität

Gisela Çakir
Dr. A.E. Çakir



ERGONOMIC INSTITUT
für Arbeits- und Sozialforschung
Berlin

1 Zielsetzung und Beschreibung

Dieser Test dient der subjektiven Beurteilung der Darstellungsgüte von Bildschirmen hinsichtlich ergonomischer Kriterien. Diese können sowohl eigenständige Bildschirmgeräte, sprich sog. Monitore, als auch Komponenten eines Arbeitsplatzrechners sein. Man kann auch Zeichensätze für eine Anwendung vergleichen.

Der Test prüft die Darstellungsgüte des Bildschirms, die während des Betriebes aufgrund verschiedener Faktoren geringer sein kann als bei einer Geräteprüfung bzw. bei der Beschaffung. Die Darstellungsgüte kann aus folgenden Gründen gemindert sein:

Ursache	Mögliche Wirkungen
Alterung der Anzeige	Abnahme der Bildschärfe, Abnahme des Kontrastes,
Verschmutzung	Abnahme der Bildschärfe, Abnahme des Kontrastes,
Einbrennen	Abnahme der Bildschärfe, Abnahme des Kontrastes, schlechte Lesbarkeit,
magn. Beeinflussung	Bildzittern, scheinbare Unschärfe des Bildes, Flimmerempfindung, bei Farbbildschirmen Konvergenz- fehler (farbige Symbole und Zei- chen wirken zerfasert).

Das Testmaterial kann erfahrungsgemäß in 20 Minuten bearbeitet werden.

Obwohl bei diesem Test die Leseleistung als objektives Gütemerkmal von Bildschirmen nicht erfaßt wird, zeigen ausländische Untersuchungen, daß das Ergebnis eines solchen Tests dem Ergebnis von objektiven Untersuchungen entspricht. Es ist jedoch kein Ersatz für objektive Tests, die bei der Geräteprüfung angewendet werden.

Dieses Testverfahren kann dazu eingesetzt werden, vorhandene Bildschirme im Betrieb zu prüfen bzw. bei Beschaffungen ansonsten gleichwertige Bildschirme hinsichtlich ergonomischer Kriterien einer vergleichenden Prüfung zu unterziehen. Durch den Test können Hinweise für die o.g. Qualitätsminderungen gewonnen werden. Das Verfahren ist kein Ersatz für Geräteprüfungen, die nur von autorisierten Stellen vorgenommen werden können. Man kann auch die Eignung von Zeichensätzen prüfen oder die Qualität gescannter Dokumente vergleichen.

2 Vorbereitung und Durchführung des Tests

2.1 Vorbereitung

1. **Vereinbarungen treffen**
Zahl der Testpersonen, Nutzung der Ergebnisse, Vertraulichkeit, Zusammensetzung der möglicherweise Beteiligten (z.B. Lieferant, Organisationsabteilung, Personalvertretung, Sicherheitsfachkraft, Betriebsarzt, externe Sachverständige)
2. **Test vorstellen**
Aufgaben, Ziele, Vordrucke, Reihenfolge der Geräte u.ä.
3. **Anonymität zusichern**
Behandlung der Ergebnisse erläutern
4. **Gerät erläutern**
Handhabung der Stellteile (z.B. Helligkeitsregelung, Ein-Aus-Schalter)
5. **Noch offene Fragen klären**

2.2 Testbedingungen

Testpersonen

Das Testkollektiv sollte in der Altersstruktur etwa der vorgesehenen Benutzerstruktur entsprechen. Die Testpersonen sollten mindestens ein Bildschirmgerät aus eigener Erfahrung kennen.

Für Orientierungstests reicht ein Kollektiv aus etwa 5 Personen. Wenn eine statistisch bedeutsame Auswertung angestrebt wird, sollte das Testkollektiv aus mindestens 10, besser aus 20 Personen bestehen.

Testraum

Der Test sollte entweder am Einsatzort, oder, wenn dies nicht möglich ist, in einem Raum mit gleichbleibenden Beleuchtungsbedingungen durchgeführt werden, die denen am Einsatzort entsprechen.

Testanordnung

Die Geräte sind unter Beobachtungsbedingungen zu prüfen, die ihrer im Betrieb üblichen Benutzung entsprechen. Es ist insbesondere wichtig, die Sehentfernung richtig einzustellen, da Bildschirme, die für große Sehentfernungen vorgesehen sind, aus geringeren Sehentfernungen i.d.R. unschärfer erscheinen. Auch die Flimmerempfindung ist unter diesen Bedingungen stärker.

Die Geräte sind so anzuordnen, wie sie in dem Betrieb üblicherweise benutzt werden bzw. in Zukunft benutzt werden sollen.

Wenn mehrere Geräte als Alternativen geprüft werden sollen, sollte das Testbild für alle Geräte gleich sein (gleiche Maske, gleiche Zahl von Schriftzeichen, Grafikelementen u.ä.)

2.3 Das Testbild

Das Testbild sollte einer typischen Abbildung (z.B. Maske oder Fließbild) aus der beabsichtigten Anwendung entsprechen, die möglichst viele unterschiedliche Zeichen und eventuelle Strichelemente aufweist. Für Bildschirme, die für Textverarbeitung eingesetzt werden, sollte der Bildschirm mit einem fortlaufenden Fließtext vollgeschrieben werden, so daß etwa 70 % aller Zeichenstellen besetzt sind. Für Bildschirme, die Fließbilder anzeigen sollen, sollten Abbildungen verwendet werden, bei denen die Strich- und die Schrift Elemente nicht nur aus den Primärfarben (Rot, Grün, Blau) bestehen, sondern möglichst viele der in der Anwendung verwendeten Mischfarben aufweisen sollten.

Das Testbild muß die vorgesehene Art der Schriftdarstellung (Groß- und Kleinschreibung oder nur Großschreibung) aufweisen. Bei Geräten, deren Zeichensatz variabel ist, sollte für den Test die Standardschrift für die jeweilige Anwendung verwendet werden.

***Anm.:** Es ist nicht sinnvoll, den Bildschirm mit gleichen Zeichen vollzuschreiben, da Testpersonen in diesem Fall ein Flimmern wahrnehmen können, das vom Muster der Zeichen herrührt und nicht von der Bildröhre. Das Testbild darf auch nicht aus gleichmäßig angeordneten geometrischen Mustern bestehen, da solche Muster u.U. selber eine Flimmerempfindung hervorrufen können.*

Wichtig ist die richtige Einstellung der Helligkeit von LCD-Monitoren, weil deren Flimmerwirkung durch Dimmen erhöht werden kann.

TEST **zur Bestimmung der** **Darstellungsgüte des Bildschirmes**

Ziel

Dieser Test dient der Beurteilung der Güte der Bildschirmdarstellung. Er besteht aus drei Untertests:

Untertest 1: Flimmerfreiheit und Stabilität

Dieser Test dient zur Feststellung der Flimmerfreiheit des Bildes und der Stabilität. Das Flimmern des Bildes und ein eventuelles Zittern der Buchstaben oder der Zeilen lassen sich zwar auf unterschiedliche technische Gründe zurückführen, ihre störende Wirkung auf den Benutzer läßt sich jedoch auf der gleichen Skala messen.

Untertest 2: Schärfe des Bildes

Dieser Test dient der Feststellung der Schärfe des Bildschirmes. Die subjektive Bewertung der Schärfe ist i.d.R. zuverlässiger als die wesentlich aufwendigere technische Messung.

Untertest 3: Schriftgüte

Dieser Test erfaßt die Beurteilung des Schriftbildes u.a. nach Gestalt sowie nach Anmutungs- und Wahrnehmungsqualität. Erfahrungsgemäß entspricht das mit diesem Test erfaßte Urteil etwa dem Ergebnis, das sich nach umfangreichen technischen Untersuchungen ergibt, die sehr zeitraubend sind, und stimmt mit dem Ergebnis von Leseleistungstests gut überein.

Bewertung

Bei zwei der drei Untertests finden Sie eine Liste von Begriffspaaren (z.B. hell - dunkel). Bitte beurteilen Sie das Gerät nach Ihrem persönlichen Eindruck und kreuzen Sie Ihr Urteil wie in dem folgenden Beispiel an:

BEISPIEL

hell ① ② ③ ④ ⑤ ⑥ dunkel

Die Bezeichnung "hell" trifft voll und ganz zu.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Bezeichnung "hell" trifft weitgehend, jedoch nicht voll und ganz zu.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Bezeichnung "hell" trifft nur teilweise zu.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Bezeichnung "dunkel" trifft eher zu als "hell".	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Bezeichnung "dunkel" trifft weitgehend zu.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Die Bezeichnung "dunkel" trifft voll und ganz zu.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Wenn Sie Fragen zur Beurteilung haben, stellen Sie diese bitte sofort!

Test zur Bestimmung der Darstellungsgüte des Bildschirms

Untertest 1: FLIMMERFREIHEIT UND STABILITÄT

Mit diesem Test möchten wir feststellen, ob dieser Bildschirm als "flimmerfrei" empfunden wird und ein stabiles Bild aufweist, wie Sie es von gedrucktem Papier her gewohnt sind.

Im folgenden finden Sie sieben Begriffspaare, die Ihrem Empfinden entsprechend angekreuzt werden sollen. Es kommt hierbei nur auf Ihre persönliche Beurteilung an. Bitte lesen Sie die Anleitung auf der nächsten Seite sorgfältig durch, bevor Sie die Bewertung vornehmen.

Bitte stellen Sie jetzt den Bildschirm so hell wie möglich ein und blicken Sie am Bildschirm seitlich vorbei. Beurteilen Sie bitte die Stabilität der Abbildung, ohne auf den Bildschirm zu schauen.

Ihr Urteil über



angenehm	①	②	③	④	⑥	⑥	unangenehm	1
stabil	①	②	③	④	⑥	⑥	zappelig	2
stehend	①	②	③	④	⑥	⑥	bewegt	3
ruhig	①	②	③	④	⑥	⑥	unruhig	4
weich	①	②	③	④	⑥	⑥	hart	5
nicht störend	①	②	③	④	⑥	⑥	störend	6
nicht flimmernd	①	②	③	④	⑥	⑥	flimmernd	7

**Test
zur Bestimmung der
Darstellungsgüte des Bildschirms**

Untertest 2: Schärfe des Bildes

Ihr Urteil über



Mit diesem Test möchten wir Ihr Urteil über die Bildschärfe des Bildschirms erfassen. Bitte beurteilen Sie den Bildschirm mit den Proben 1 bis 6 (s. nächste Seite) und kreuzen Sie die Nummer der Probe an, die der Qualität des Bildes auf dem Bildschirm am ähnlichsten erscheint.

Urteil 1: Beurteilung der grafischen Objekte

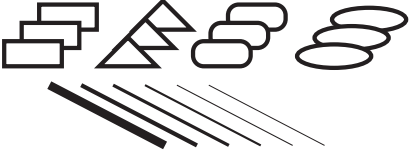
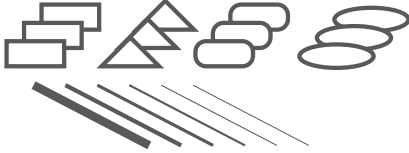




Welcher Probe sieht das Bild am ähnlichsten?:

Urteil ① ② ③ ④ ⑤ ⑥

Urteil 2: Beurteilung der Schrift

Welcher Schriftprobe sieht das Schriftbild am ähnlichsten:

Urteil ① ② ③ ④ ⑤ ⑥

 <p>Die Beschriftung von grafischen Abbildungen muß klar und deutlich erkennbar sein.</p> <p>①</p>	 <p>Die Beschriftung von grafischen Abbildungen muß klar und deutlich erkennbar sein.</p> <p>i</p>
 <p>Die Beschriftung von grafischen Abbildungen muß klar und deutlich erkennbar sein.</p> <p>③</p>	 <p>Die Beschriftung von grafischen Abbildungen muß klar und deutlich erkennbar sein.</p> <p>④</p>
 <p>Die Beschriftung von grafischen Abbildungen muß klar und deutlich erkennbar sein.</p> <p>⑤</p>	 <p>Die Beschriftung von grafischen Abbildungen muß klar und deutlich erkennbar sein.</p> <p>⑥</p>

Test zur Bestimmung der Darstellungsgüte des Bildschirms

Untertest 3: Schriftgüte

Ziel

Dieser Test erfaßt die Beurteilung des Schriftbildes nach Gestalt der Zeichen, Anmutungsqualität des Schriftbildes, Zeilenabstand, Schriftproportionen und Wahrnehmungsqualität. Mit dem Test ist es möglich, umfangreiche technische Untersuchungen, wie Leuchtdichte- und Kontrastmessungen, Strichstärkebestimmung u.ä., näherungsweise durch einen einfachen Test zu ersetzen, der im Betrieb durchgeführt werden kann.

Faktoren

Faktor 1	QUALITÄT Beurteilung der Zeichengestaltung, Zeichenqualität
Faktor 2	ÄSTHETIK Beurteilung der Anmutungsqualität des Schriftbildes
Faktor 3	STÄRKE Beurteilung der Form der Zeichen und des Schriftbildes auf der Dimension "hart, gezackt – weich, fein"
Faktor 4	WAHRNEHMUNGSQUALITÄT Beurteilung der Schärfe und der Deutlichkeit des Schriftbildes

Test
zur Bestimmung der
Darstellungsgüte des Bildschirms

Untertest 3: Schriftgüte

Ihr Urteil über



Mit diesem Test möchten wir Ihr Urteil über die Qualität der Schrift auf diesem Bildschirm erfassen. Es gibt bekanntlich viele Schriften, von denen wir einige unten beispielhaft zeigen. Sie sind alle unterschiedlich in der Gestalt und in der Lesbarkeit.

Ein Anfang ist
kein Meisterstück,
doch ein guter Anfang
halbes Glück

Ein Anfang ist
kein Meisterstück,
doch ein guter Anfang
halbes Glück

Ein Anfang ist
kein Meisterstück,
doch ein guter Anfang
halbes Glück

Ein Anfang ist
kein Meisterstück,
doch ein guter Anfang
halbes Glück

Ein Anfang ist
kein Meisterstück,
doch ein guter Anfang
halbes Glück

**Ein Anfang ist
kein Meisterstück,
doch ein guter Anfang
halbes Glück**

Stellen Sie bitte die Helligkeit des Bildschirmes so ein, daß Sie den darin dargestellten Text gut lesen können. Bitte lesen Sie den Inhalt des Bildschirmes in aller Ruhe durch.

Jetzt beurteilen Sie das Schriftbild in derselben Weise, wie Sie das Flimmern beurteilt haben. Diesmal sollen Sie 23 Begriffspaare wie "gut lesbar – schlecht lesbar" oder "harmonisch – unharmonisch" benutzen. Entscheidend ist wieder Ihr persönlicher Eindruck und Ihre Vorstellung von einer gut lesbaren Schrift.

**Test
zur Bestimmung der
Darstellungsgüte des Bildschirms**

Untertest 3: Schriftgüte

Ihr Urteil über



gut lesbar	①	②	③	④	⑤	⑥	schlecht lesbar	8
scharf	①	②	③	④	⑤	⑥	unscharf	9
weich	①	②	③	④	⑤	⑥	hart	10
klar	①	②	③	④	⑤	⑥	unklar	11
gut gestaltet	①	②	③	④	⑤	⑥	schlecht gestaltet	12
akzeptabel	①	②	③	④	⑤	⑥	nicht akzeptabel	13
ausgewogen	①	②	③	④	⑤	⑥	unausgewogen	14
hell	①	②	③	④	⑤	⑥	dunkel	15
deutlich	①	②	③	④	⑤	⑥	undeutlich	16
stark	①	②	③	④	⑤	⑥	schwach	17
fein	①	②	③	④	⑤	⑥	gezackt	18
angenehm	①	②	③	④	⑤	⑥	unangenehm	19
zumutbar	①	②	③	④	⑤	⑥	unzumutbar	20
interessant	①	②	③	④	⑤	⑥	langweilig	21
komfortabel	①	②	③	④	⑤	⑥	unkomfortabel	22
gut erkennbar	①	②	③	④	⑤	⑥	schlecht erkennbar	23
harmonisch	①	②	③	④	⑤	⑥	unharmonisch	24
leicht	①	②	③	④	⑤	⑥	schwierig	25

3 Auswertung

Die Auswertung von Testdaten erfolgt üblicherweise mit Hilfe von statistischen Verfahren. Nachfolgend wird eine vereinfachte Auswertung und Bewertung beschrieben, die für den vorgesehenen Zweck hinreichen.

3.1 Information verdichten

Das Ergebnis der Untersuchung kann in unterschiedlicher Art und Weise aufbereitet werden. Statistische Verfahren zur Auswertung der Daten sollten erst ab einer Größe des Testkollektivs von 10 Personen angewendet werden. Für betriebliche Untersuchungen genügt eine grafische Darstellung von Mittelwerten. Zu diesem Zweck wird für jede einzelne Bewertung der Mittelwert über alle Testpersonen gebildet.

3.2 Profildarstellung

Um die Profildarstellung (s. unten) zu erhalten, wird zu jedem Testpunkt der Mittelwert der Urteile des Testkollektivs gebildet. Man kann zu jedem Untertest ein Profil zeichnen oder die Daten weiter verdichten, indem man die Urteile auf bestimmten Skalen zusammenfaßt. Das unten angeführte Beispiel zeigt Beurteilungsprofile von zwei Bildschirmen. Aufgrund früher durchgeführter Untersuchungen ist folgende Zusammenfassung sinnvoll:

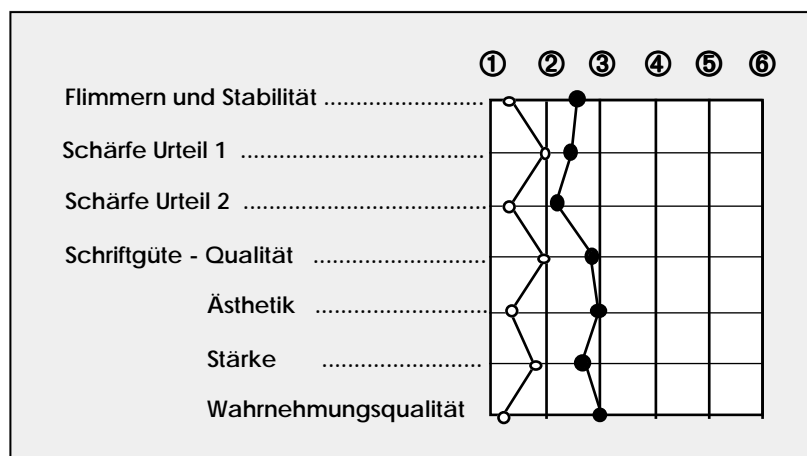
Flimmern und Stabilität Mittelwert von 1 bis 7

Schärfe des Bildes

Urteil 1 Mittelwert von Untertest 2, Urteil 1
 Urteil 2 Mittelwert von Untertest 2, Urteil 2

Schriftgüte

Qualität Mittelwert von 8, 13, 20, 22, 23, 25
 Ästhetik Mittelwert von 14, 19, 21, 24
 Stärke Mittelwert von 10, 15, 17, 18
 Wahrnehmungsqualität Mittelwert von 9, 11, 16



3.3 Bewertung und Nutzung des Ergebnisses

Die besten der heute verfügbaren Bildschirme erreichen auf allen Faktoren Werte, die sich nur geringfügig von "1" unterscheiden. Bei kleinen Testkollektiven liegt der Erwartungswert für einwandfreie Bildschirme zwischen "1" und "2".

Wenn ein Bildschirm auf einem oder mehr Faktoren vom Erwartungswert abweicht, sollte das betreffende Gerät näher untersucht werden. Sofern eine Abweichung bei "Flimmern und Stabilität" festgestellt wird, sollten die möglichen Abhilfen aus Anleitung 7 in Erwägung gezogen werden. Wenn das Urteil über die "Schärfe des Bildes" negativ ausfällt, sollte zunächst geprüft werden, ob die Bildfläche verschmutzt ist (*wahrscheinlichste Ursache, wenn das Gerät im Originalzustand ein scharfes Bild liefert*). Ist dies nicht der Fall, sollte die Bildeinstellung geprüft werden.

Wenn "Flimmern und Stabilität" und "Schärfe des Bildes" gleichzeitig negativ beurteilt werden, kann die Ursache in einer Beeinflussung durch elektromagnetische Felder bestehen. Wenn dies zutrifft, sollte das Bild des Geräts noch einmal an einem anderen Standort geprüft werden. (*Anm.: Wenn die hier genannten Faktoren negativ beurteilt werden, wird die Schriftgüte i.d.R. auch ungünstig beurteilt.*)

Wird "Flimmern und Stabilität" und "Schärfe des Bildes" günstig beurteilt (Werte zwischen "1" und "2"), jedoch nicht die "Schriftgüte", liegt das Problem vermutlich beim Zeichensatz der Anwendung. Dieser kann bei Arbeitsplatzcomputern i.d.R. geändert werden, jedoch nicht bei Terminals von Prozeßrechnern und Hostrechnern.

Wird "Flimmern und Stabilität" günstig beurteilt (Werte zwischen "1" und "2"), jedoch nicht "Schärfe des Bildes" und "Schriftgüte", kann die Ursache durch Einbrennen oder allgemeines Altern des Monitors entstanden sein.

Vergleichende Bewertungsprofile (s. S. 13) können nur für vergleichbare Bildschirme (z.B. Farbmonitore, monochrome Anzeigen) sinnvoll interpretiert werden. Zwei Bildschirme sind bezüglich eines Kriteriums (z.B. "Flimmern und Stabilität") nur dann als unterschiedlich bewertet anzunehmen, wenn die Bewertungen sich um etwa einen Punkt auf der jeweiligen Skala unterscheiden. Eine statistisch korrekte Bestimmung der Unterschiedlichkeit läßt sich mit dem t-Test für abhängige Stichproben vornehmen.