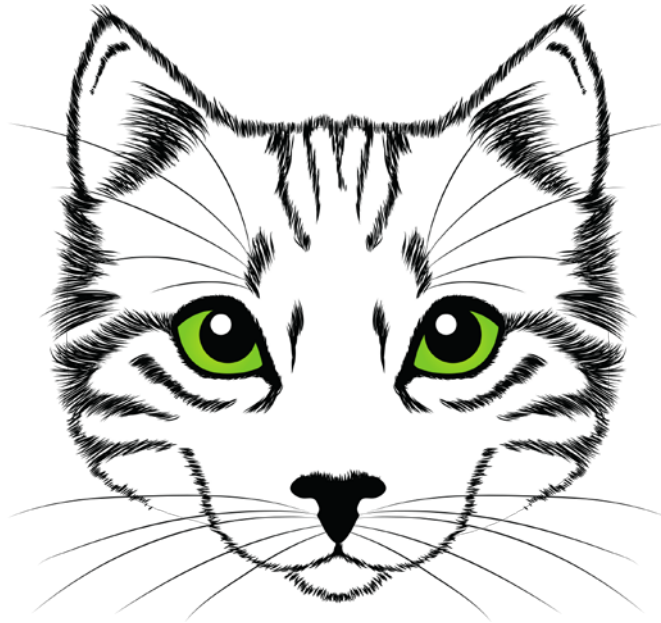


Klausurerstellung und -auswertung an der Fakultät 11 der LMU



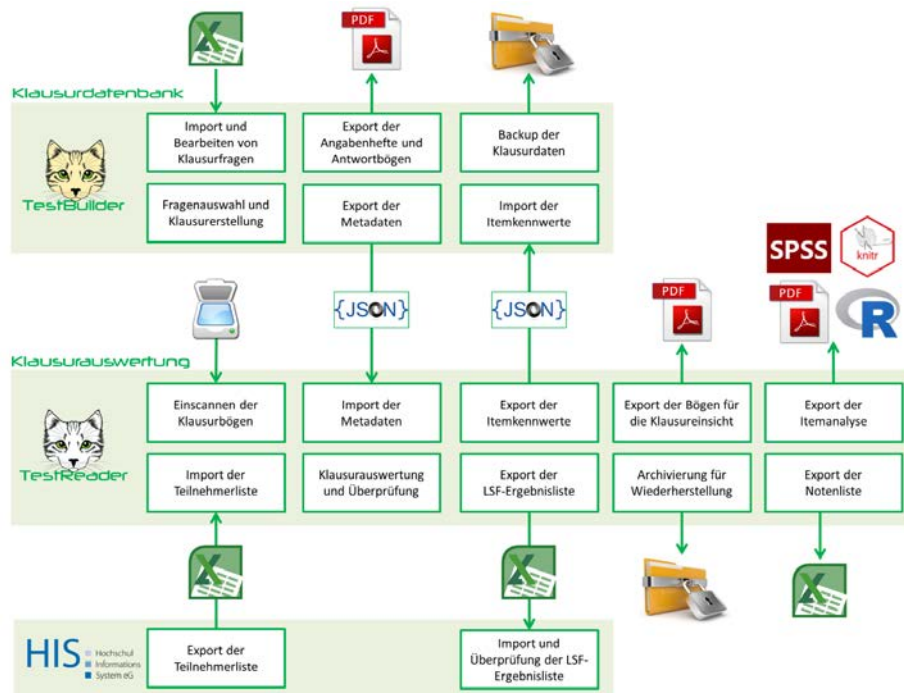
spe TestBuilder/TestReader

Die Programme **TestBuilder** und **TestReader** des **Lehrstuhls für Schulpädagogik** (= spe) sind Computeranwendungen zur einfachen Erstellung sowie schnellen und sicheren Auswertung von Multiple-Choice-Klausuren. Sie werden seit 2010 am Lehrstuhl für Schulpädagogik der LMU München von Dr. Thomas Lerche fortlaufend entwickelt und werden mittlerweile an der Fakultät 11 der LMU für die Erstellung und Auswertung von Multiple-Choice-Klausuren eingesetzt. Die Programme führen nicht nur sämtliche notwendigen Arbeitsschritte aus, sondern bieten in Funktionalität und Bedienbarkeit zahlreiche Vorteile gegenüber herkömmlichen, proprietären Softwarelösungen.

Wie funktionieren die Programme?

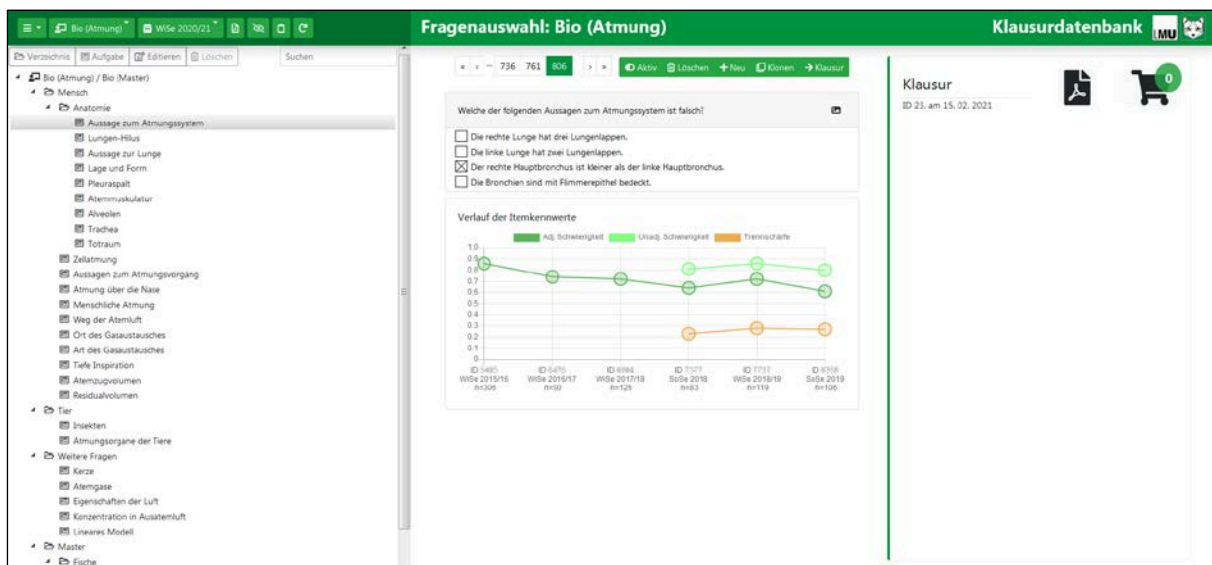
TestBuilder und TestReader ergänzen einander. Der TestBuilder verwaltet den Fragenkatalog und ermöglicht ein einfaches und schnelles Zusammenstellen der Klausur in verschiedenen Versionen. Der TestReader liest die Klausurscans ein, wertet diese aus und exportiert die HIS-Notenlisten sowie die Analyse der Klausurfragen hinsichtlich Schwierigkeiten und Trennschärfe.

Organisatorisch läuft die Erstellung und Auswertung einer Klausur mit Hilfe von TestBuilder und TestReader folgendermaßen ab:



TestBuilder

Der TestBuilder verwaltet den Fragenkatalog. Er dient zum Sammeln der Items und zum Erstellen der Klausuren. Dank der übersichtlichen Struktur des Programmes stellt sich die Bedienung sehr intuitiv dar, wodurch er ohne große Vorkenntnisse genutzt werden kann. Auch bietet er die Möglichkeit, Fragenkataloge sowie Prüfungen für verschiedene Veranstaltungen anzulegen und zu verwalten.



Im TestBuilder werden Klausurfragen erstellt bzw. importiert, systematisch geordnet und im Fragenkatalog archiviert. Dieser kann einerseits neu zusammengestellt werden und es können jederzeit manuell Fragen hinzugefügt, bearbeitet oder gelöscht werden. Er kann ebenso aus einer vorhandenen Datenbank (z.B. von der Vorgänger-Software) aufgebaut werden.

Die angenehmste Art der Erweiterung des Fragenkatalogs besteht aus dem Import einer speziell dafür angelegten Excel-Datei. Diese wird vom Dozierenden der Veranstaltung mit neuen Fragen und den dazugehörigen Antworten versehen. Daraus werden diese dann mit ein paar Mausklicks in den Fragenkatalog importiert. Bei Bedarf können die gewünschten Fragen hier auch direkt in die Klausur des aktuellen Semesters übernommen werden.

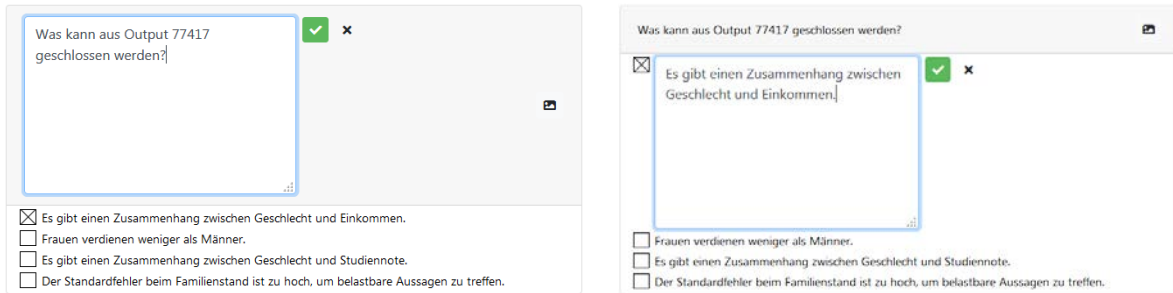
The screenshot shows the 'Excel-Import: Bio (Atmung)' interface. The left pane lists various questions related to respiration. The central area shows the 'Importieren' button and a preview of the imported data. The right pane shows a hierarchical tree of subjects and topics, including 'Bio (Atmung) / Bio (Master)', 'Mensch', 'Anatomie', 'Zellatmung', 'Aussagen zum Atmungsvorgang', 'Atmung über die Nase', 'Menschliche Atmung', 'Weg der Atemluft', 'Ort des Gasaustausches', 'Art des Gasaustausches', 'Tiefe Inspiration', 'Atemzugvolumen', 'Residualvolumen', 'Tier', 'Insekten', 'Atmungsorgane der Tiere', 'Weitere Fragen', 'Kerze', 'Atemgase', 'Eigenschaften der Luft', 'Konzentration in Ausatemluft', 'Lineares Modell', 'Master', 'Fische', 'Körperbau 1', 'Körperbau 2', 'Körperbau 3', and '123'. The bottom of the right pane shows '6 Aufgaben'.

Jede dieser Fragen kann vor Abschluss der Klausurerstellung bearbeitet werden. Dies gilt auch für bereits vorhandene Fragen aus vorangegangenen Prüfungen. Jedoch werden diese zuvor automatisch geklont, da sie in ihrem Urzustand auf Grund der Archivierungssicherheit nicht verändert werden dürfen. Dieser Vorgang ermöglicht daher trotzdem das Verwenden und Überarbeiten bewährter Fragen.

The screenshot shows the question management interface. At the top, a green notification bar states: 'Die Frage 711 wurde als Frage 801 geklont.' Below this is a navigation bar with buttons for 'Aktiv', 'Löschen', '+ Neu', 'Klonen', and '→ Klausur'. The main area displays a question card with the text: 'Was kann aus Output 77417 geschlossen werden?' and four multiple-choice options:

- Es gibt einen Zusammenhang zwischen Geschlecht und Einkommen.
- Frauen verdienen weniger als Männer.
- Es gibt einen Zusammenhang zwischen Geschlecht und Studiennote.
- Der Standardfehler beim Familienstand ist zu hoch, um belastbare Aussagen zu treffen.

Wurde eine Frage noch nicht verwendet bzw. wie angesprochen geklont, kann sie direkt auf der Darstellungssseite bearbeitet werden. Ein einfacher Klick auf den Fragentext öffnet den integrierten Inline-Editor, dargestellt durch ein Textfeld. Änderungen werden somit unmittelbar wirksam. Dies gilt auch für die einzelnen Antwortmöglichkeiten oder die Markierungen (Richtig/Falsch).



Bilder und Grafiken werden im Editor direkt eingefügt. Das System unterstützt gängige Dateiformate wie .png oder .jpg.

711

Aktiv Löschen + Neu Klonen → Klausur

Was kann aus Output 77417 geschlossen werden?

```

*****
***output_77417***
*****

. * Lineare Regression AV=Monatliches Einkommen erklärt durch
Geschlecht (0=weiblich, 1=männlich), Studiennote (Noten: 1-4),
Familienstand (0=ledig, 1=verheiratet)

. reg Einkommen Geschlecht Studiennote Famstand, beta

-----
Source |      SS      df    MS    Number of obs   =    8,573
-----+-----
Model   | 1481316.38    3  493772.127   F(3, 8569)       =   10.57
Residual| 400243368    8,569 46708.2936   Sig. (F)         =   0.0000
-----+-----
Total   | 401724684    8,572 46864.7555   R2              =   0.0372
-----+-----
                Korrr. R2    =   0.0330
                Root MSE     =   216.12

-----
Einkommen |      b      Std. Err.      t    P>|t|      Beta
-----+-----
Geschlecht | 23.61074   4.832602    4.89  0.000   .0532551
Studiennote | 7.250289   3.888036    1.86  0.062   .0203339
Famstand   | 14.0421    13.80163    1.02  0.309   .010978
b0         | 1554.655   9.038672   172.00 0.000
-----+-----

. vif

-----
Variable |      VIF      Tol
-----+-----
Studiennote | 1.02  0.977852
Geschlecht | 1.02  0.978588
Famstand   | 1.00  0.998671
-----+-----
Mean VIF | 1.02
                
```

(130,36 KB; image/png)

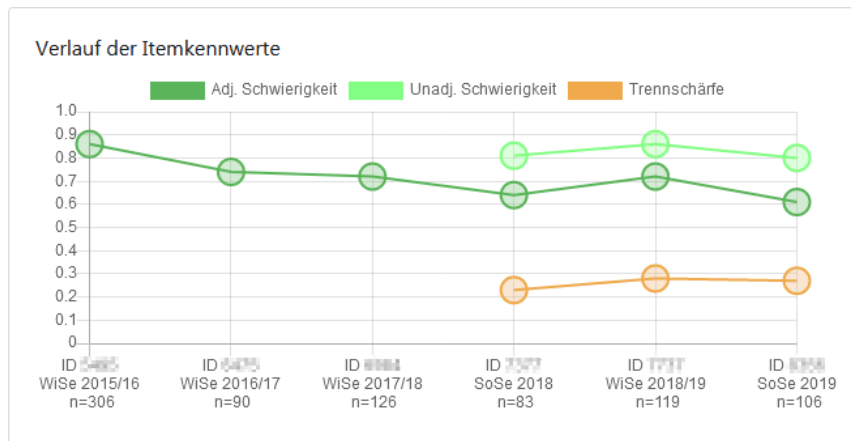
Es gibt einen Zusammenhang zwischen Geschlecht und Einkommen.

Frauen verdienen weniger als Männer.

Es gibt einen Zusammenhang zwischen Geschlecht und Studiennote.

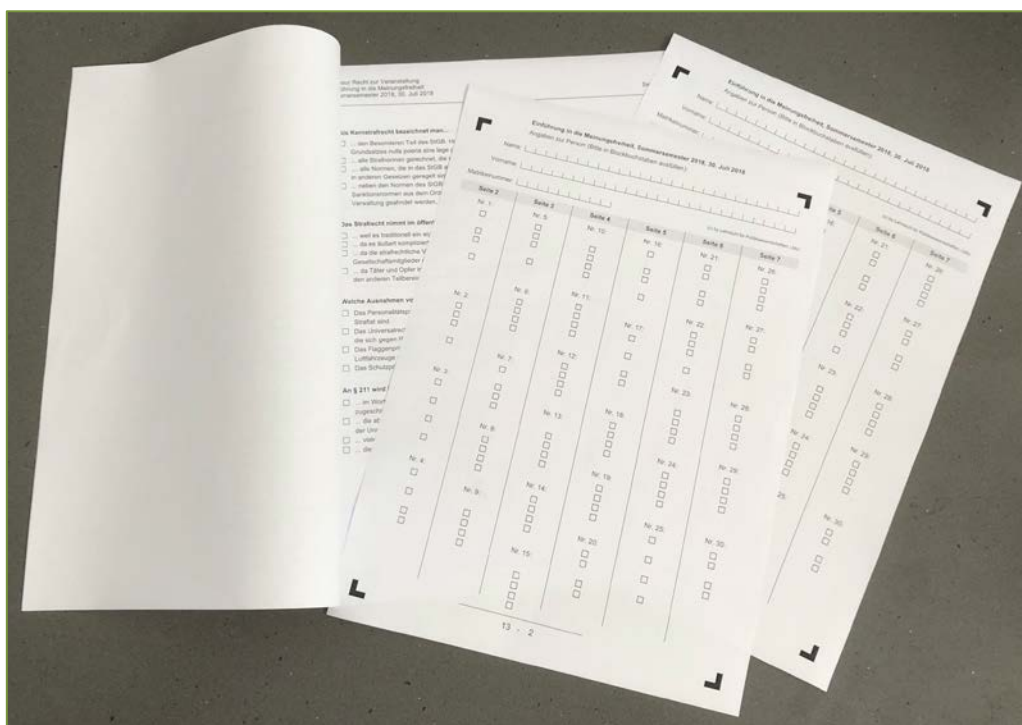
Der Standardfehler beim Familienstand ist zu hoch, um belastbare Aussagen zu treffen.

Das System präsentiert zu jeder Frage den Verlauf der Itemkennwerte.



Aus dem angesprochenen Fragenkatalog können nun einzelne Fragen per Mausklick selektiert und mit wenigen Klicks in die Klausur übernommen werden. Der aktuelle Stand der Klausur kann jederzeit über eine Vorschaufunktion eingesehen werden.

Sind alle Fragen für die aktuelle Klausur ausgewählt und kontrolliert, bietet der TestBuilder noch die Möglichkeit, vor Abschluss der Klausurerstellung eine beliebige Anzahl verschiedener Varianten (d.h. Versionen der Prüfung mit randomisierter Reihenfolge der Fragen und/oder Kategorien) zu erstellen. Die Ausgabe der Prüfung erfolgt hierbei im allgemein weit verbreiteten PDF-Format. Der Druck der Prüfungen erfolgt über die Hausdruckerei. Sie erstellt pro Prüfungsteilnehmer jeweils ein Aufgabenheft und zwei sogenannte Antwortbögen. Nur die letztgenannten werden ausgewertet.



TestReader

Nachdem die Antwortbögen von den Prüflingen ausgefüllt wurden, werden diese vor dem Einscannen zunächst von Hand auf Verwertbarkeit geprüft und gezählt. Das anschließende Einscannen mit Hilfe eines handelsüblichen Scanners läuft wie gewohnt ab.

Nun beginnt die Arbeit des TestReaders: Er erfasst nicht nur, ob die einzelnen Kästchen vom Prüfling angekreuzt wurden oder nicht, sondern wertet auch die handschriftlichen deskriptiven Angaben aus und weist den Antwortbogen so dem entsprechenden Prüfling zu. Diese Erkennung läuft automatisiert ab.

The screenshot shows the 'spe TestReader 0.9.9.2 - Klausur id7' application. The left pane displays a list of scanned pages with columns for 'Nr.', 'MatNr.', 'Datei', 'Skal.', 'Vers.', 'Ausg.', and 'Zugr.'. The right pane shows the scanned content of page 13, which is a form titled 'Einführung in die Schulpädagogik, Sommersemester 2018, 10. Juli 2018'. The form includes fields for 'Name', 'Vorname', and 'Matrikelnummer'. Below these fields is a grid of 40 numbered items (Nr. 1 to Nr. 40) arranged in 10 columns (Seite 2 to Seite 10) and 4 rows. Each item has a small box with a '0' or '1' next to it, indicating whether it was marked. The status bar at the bottom shows '851/851 skaliert', '851/851 ausgewertet', and '851/851 zugewiesen'.

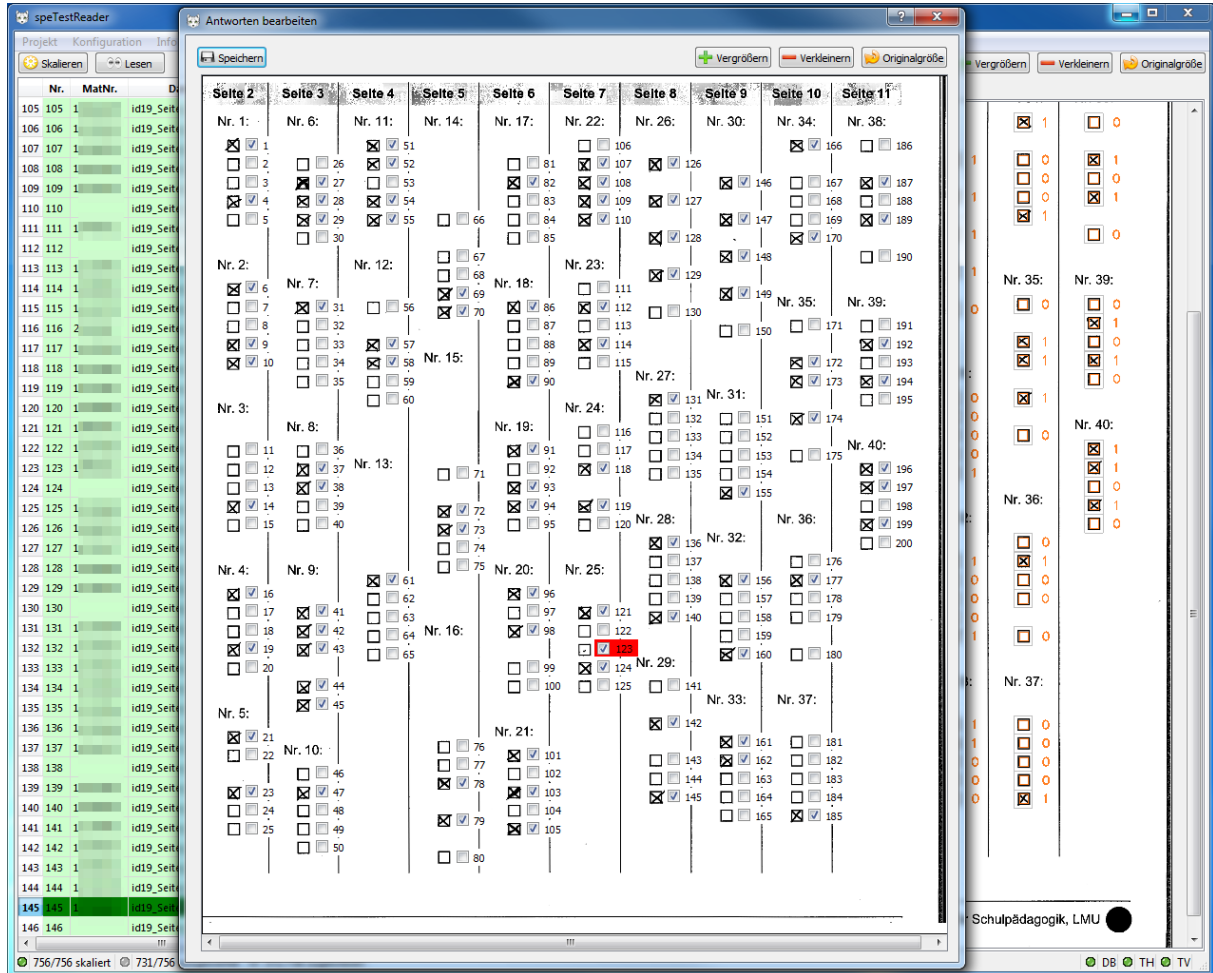
Sollten die Angaben aufgrund eines unreinen Schriftbildes vom TestReader nicht eindeutig zugeordnet werden können, werden die entsprechenden Informationen auf dem Bildschirm angezeigt. Zusätzlich wird ein Vorschlag für den wahrscheinlichsten Namen präsentiert. Dieser braucht in der Regel nur noch mit einem Druck der Enter-Taste bestätigt werden.

The screenshot shows the 'spe TestReader 0.9.9.2 - Klausur id7' window. The interface is divided into several sections:

- Project List (Left):** A table with columns for 'Nr.', 'MatrNr.', 'Datei', 'Skal.', 'Vers.', 'Ausg.', and 'Zugr.'. It lists 42 document pages from 'id7_Seite_0022.png' to 'id7_Seite_0063.png'. The 61st and 62nd rows are highlighted in green.
- Document Preview (Center):** Shows the original scan and recognized answers for three pages. The first two pages contain names 'P. Pfeiffer' and 'J. Pfeiffer' with redaction boxes. The third page shows handwritten numbers '1120046?' and '1120046?'. Below the preview, the name 'P. Pfeiffer' is suggested, followed by 'J. Pfeiffer' and '1120046?'.
- Evaluation Table (Right):** A table titled 'Matrikelnummer (Name) Sicherheit' listing 20 rows with student IDs (e.g., 1130601) and scores (e.g., 7,5). The first row is highlighted in green.

At the bottom of the window, there are status indicators for '851/851 skaliert', '851/851 ausgewertet', '850/851 zugewiesen', and a row of icons (DB, TH, TV).

Die Testantworten erkennt der TestReader nahezu perfekt. Falls ein Kästchen aber nicht sicher gewertet werden kann, bedarf dieses einer zusätzlichen Inaugenscheinnahme. Auch diese manuelle Überprüfung wird vom TestReader unterstützt: Das Programm setzt an der entsprechenden Stelle eine Markierung. In einem einfach zu bedienenden Dialog kann der Vorschlag des TestReaders dann bestätigt oder korrigiert werden.



Abschließend muss noch überprüft werden, ob der TestReader alle Antworten richtig erkannt und alle deskriptiven Daten korrekt zugeordnet hat. Diese prüfungsrechtlich gebotene Aufgabe wird vom TestReader ebenfalls erleichtert. Das Programm zeigt alle Ergebnisse eines Testbogens auf einer Seite an. Nach der Überprüfung der deskriptiven Daten können die erkannten Antworten in Sekundenschnelle überprüft werden: Alle Kästchen, die vom TestReader als "angekreuzt" erkannt wurden, werden auf der linken Seite dargestellt, alle Kästchen, die als "nicht angekreuzt" gewertet wurden, findet man auf der rechten Seite. Eventuelle Fehler können so unmittelbar entdeckt werden.

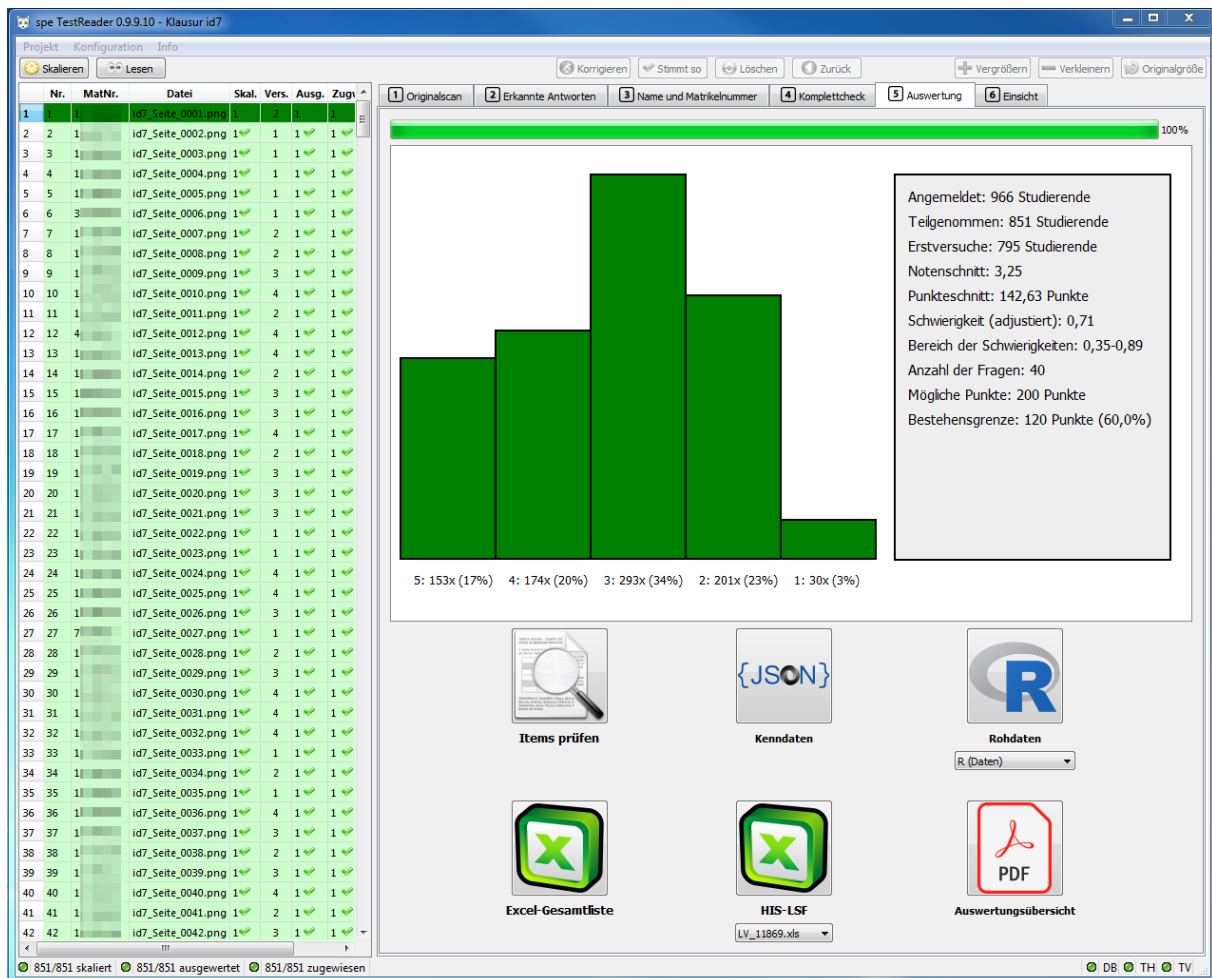
The screenshot displays the 'spe TestReader 0.9.9.2 - Klausur id7' application. The interface is divided into several sections:

- Project Information:** 'Projekt: id7', 'Konfiguration', 'Info'.
- Navigation:** Buttons for 'Skalieren', 'Lesen', 'Korrigieren', 'Stamm so', 'Berichtigen', 'Zurück', 'PDF', 'Vergrößern', 'Verkleinern', 'Originalgröße'.
- Page List:** A table with columns: Nr., MatNr., Datei, Skal., Vers., Ausg., Zugr. The list shows pages from id7_Seite_0810.png to id7_Seite_0851.png, with status indicators (checkmarks) for each.
- View Modes:** A tabbed interface with options: 1 Originalscan, 2 Erkannte Antworten, 3 Name und Matrikelnummer, 4 Kompletcheck, 5 Auswertung, 6 Einsicht.
- Main Content Area:** Displays the scanned content of the selected page (id7_Seite_0849.png). It includes a header with 'S', 'S', 'D' and a large grid of checkboxes. The left column of checkboxes is mostly checked (marked with an 'X'), while the right column is mostly empty.
- Status Bar:** Shows '851/851 skaliert', '851/851 ausgewertet', '851/851 zugewiesen' and system icons for 'DB', 'TH', 'TV'.

Wenn alle Kontrollen und Zuordnungen durchgeführt wurden, kann die abschließende Klausurauswertung durchgeführt werden.

Das Programm berechnet dabei automatisch die Bestehensgrenze gemäß den Vorgaben der Studienordnungen der LMU München ("flexible Bestehensgrenze").

Dabei werden die Noten der einzelnen Kommilitonen stets zur Programmlaufzeit berechnet, damit keine Bewertungen in den lokalen Programmdateien gespeichert werden müssen. Die Software operiert dabei lediglich mit den Excel-Dateien, die als Listenexport von HIS-LSF erzeugt werden sowie den erkannten Punkten und Matrikelnummern der eingescannten Bögen. Zur Laufzeit werden beide Datenstrukturen über die eindeutige Variable "Matrikelnummer" zusammengeführt und die Ergebnisdateien generiert.



Folgende Dateien werden unter anderem erzeugt:

Die Excel-Datei(en) für den Import der Noten in HIS-LSF.

nr	abschl	sig	vert	schwp	version	pr	sortname	geschl	hw	partngb	pstatus	pversuch	pvermer	prsem	er
5	1	32	250	---	P2	2010	10299	A	M	SNE	LMU	AN	3	20181	A
6	1	31	250	---	P2	2016	10299	A	W	---	LMU	AN	1	20181	A
7	1	55	250	---	P1	2010	10299	A	M	---	LMU	AN	1	20181	F
8	1	35	250	---	P1	2016	10299	A	W	---	LMU	AN	2	20181	F
9	1	35	250	---	P1	2016	10299	A	W	---	LMU	AN	1	20181	A
10	1	35	250	---	P1	2016	10299	A	W	---	LMU	AN	1	20181	A
11	1	35	250	---	P1	2016	10299	A	W	---	LMU	AN	1	20181	A
12	1	35	250	---	P1	2010	10299	A	M	---	LMU	AN	1	20181	V
13	1	31	250	---	P2	2016	10299	A	W	---	LMU	AN	1	20181	C
14	1	53	250	---	P1	2016	10299	A	M	SNE	LMU	AN	1	20181	C
15	1	35	250	---	P1	2016	10299	A	W	---	LMU	AN	1	20181	F
16	1	38	250	---	P2	2016	10299	A	W	---	LMU	AN	1	20181	T
17	1	33	250	---	P1	2016	10299	A	W	---	LMU	AN	1	20181	T
18	1	53	250	---	P1	2016	10299	A	W	---	LMU	AN	3	20181	C
19	1	35	250	---	P1	2010	10299	A	M	---	LMU	AN	1	20181	F
20	1	35	250	---	P1	2016	10299	A	W	---	LMU	AN	1	20181	F
21	1	35	250	---	P1	2010	10299	A	M	---	LMU	AN	1	20181	F
22	1	35	250	---	P1	2016	10299	A	W	---	LMU	AN	1	20181	F
23	1	52	250	---	P2	2016	10299	A	W	---	LMU	AN	1	20181	F
24	1	35	250	---	P1	2016	10299	A	M	---	LMU	AN	1	20181	F
25	1	33	250	---	P1	2016	10299	A	W	---	LMU	AN	2	20181	F
26	1	35	250	---	P1	2016	10299	A	W	SNE	LMU	AN	2	20181	F
27	1	35	250	---	P1	2016	10299	A	W	---	LMU	AN	1	20181	S
28	1	33	250	---	P1	2016	10299	A	M	---	LMU	AN	1	20181	S
29	1	32	250	---	P2	2016	10299	A	W	---	LMU	AN	1	20181	F
30	1	35	250	---	P1	2016	10299	A	M	SNE	LMU	AN	2	20181	F
31	1	38	250	---	P2	2016	10299	A	W	---	LMU	AN	1	20181	F
32	1	58	250	---	P2	2016	10299	A	W	---	LMU	AN	1	20181	C
33	1	58	250	---	P2	2016	10299	A	W	---	LMU	AN	2	20181	C
34	1	35	250	---	P1	2010	10299	A	M	---	LMU	AN	1	20181	S
35	1	31	250	---	P2	2016	10299	A	W	---	LMU	AN	1	20181	S
36	1	35	250	---	P1	2016	10299	A	W	---	LMU	AN	1	20181	S
37	1	35	250	---	P1	2016	10299	A	W	SNE	LMU	AN	1	20181	E
38	1	51	250	---	P2	2016	10299	B	W	---	LMU	AN	1	20181	I
39	1	51	250	---	P2	2016	10299	B	W	---	LMU	AN	1	20181	I
40	1	53	250	---	P1	2016	10299	B	M	---	LMU	AN	1	20181	I
41	1	35	250	---	P1	2016	10299	B	M	---	LMU	AN	1	20181	I
42	1	35	250	---	P1	2016	10299	B	M	---	LMU	AN	1	20181	I

Eine PDF-Datei mit den Analysedaten der Klausur und der einzelnen Fragen

Klausur zur Veranstaltung "Einführung in die Schulpädagogik", Prof. Dr. Ewald Kiel, Lehrstuhl für Schulpädagogik, LMU Seite 2

1 Wodurch ist [redacted] gekennzeichnet?
 • d_adj=0.54 • d_unadj=0.76 • r=0.25
 • r=0.46 (74) Durch [redacted]
 • r=0.40 (811) Durch [redacted]
 • r=0.47 (264) Durch [redacted]
 • r=0.62 (202) Durch [redacted]
 • r=0.47 (121) Durch [redacted]

2 Wodurch lässt sich [redacted] charakterisieren?
 • d_adj=0.79 • d_unadj=0.89 • r=0.38
 • r=0.70 (621) Das [redacted]
 • r=0.34 (829) Das [redacted]
 • r=0.42 (60) Das [redacted]
 • r=0.48 (63) Das [redacted]
 • r=0.56 (87) Das [redacted]

3 Mit welchen Argumenten lässt sich [redacted] widerlegen?
 • d_adj=0.61 • d_unadj=0.80 • r=0.28
 • r=0.51 (514) Empirische Studien haben gezeigt, [redacted]
 • r=0.25 (23) [redacted]
 • r=0.58 (230) [redacted]
 • r=0.49 (679) [redacted]
 • r=0.45 (109) [redacted]

Der TestReader bereitet auch die Bögen für die Klausureinsicht vor und bietet einen Export als PDF-Datei an. Hierfür werden die eingescannten Bögen mit zusätzlichen Informationen ergänzt: Korrekte Antworten, Einzelpunkte, Gesamtpunkte, Aggregationsverlauf, Note und Rang. Diese Bögen können im Rahmen der Einsichtnahme an die Studierenden ausgegeben werden.

Zur Vorbereitung der Klausureinsicht gibt es die Möglichkeit, die Klausuren der hierfür angemeldeten Kommilitonen mittels einer Suchanfrage (z.B. eine Liste von Matrikelnummern oder Namen) zu filtern und so die gewünschten Einsichtsbögen schnell zu finden und – z. B. als PDF-Datei für den Ausdruck – zu erstellen.

The screenshot displays the 'spe TestReader 0.9.9.10 - Klausur id7' application. The interface is divided into several sections:

- Left Panel:** A list of scanned pages with columns for 'Nr.', 'Matr.', 'Datei', 'Skal.', 'Vers.', 'Ausg.', and 'Zugr.'. The list shows pages 540 through 581, all with a status of '1'.
- Top Bar:** Contains navigation and action buttons: 'Skalieren', 'Lesen', 'Korrigieren', 'Stimmt so', 'Löschen', 'Zurück', 'Vergrößern', 'Verkleinern', and 'Originalgröße'.
- Main Area:** Displays the title 'Einführung in die Schulpädagogik, Sommersemester 2016, 10. Juli 2018' and 'Angaben zur Person (Bitte in Blockbuchstaben ausfüllen):'. Below this is a grid of answer sheets for 40 students (Nr. 1 to 40) across pages 2 to 10. Each cell in the grid contains a small grid of checkboxes and numerical values, representing the student's answers and scores for each question.
- Right Panel:** A search bar labeled 'Name (Matrikel)' with a list of names below it. The names are partially obscured by a green highlight.
- Bottom Bar:** Shows status information: '851/851 skaliert', '851/851 ausgewertet', '851/851 zugewiesen', and 'DB TH TV'.

Welche Vorteile bieten die Programme?

Die Motivation hinter der Entwicklung des TestBuilders und TestReaders war die Herausforderung, die im Rahmen des erziehungswissenschaftlichen Studiums und der Bachelor/Masterstudiengänge erforderlichen Multiple-Choice-Klausuren mit ökonomischem Arbeitsaufwand sicher, elegant und fehlerfrei erstellen und auswerten zu können. Dieses Ziel wurde erreicht: Seit dem Wintersemester 2015/16 werden jedes Semester ca. 30 Klausuren mit ca. 8.000 Teilnehmer/innen (= fast 1 Mio. Kästchen) ausgewertet – dank der Software und des klaren organisationalen Prozesses ohne fehlerhafte Auswertungen.

Auch handschriftliche Zeichen erkennt die Software mit einer Genauigkeit von fast 97%, was im Bereich der marktüblichen Algorithmen für Handschriftenerkennung liegt. Bei neunstelligen Matrikelnummern werden damit ca. 75% aller eingescannten Matrikelnummern unmittelbar einer/m Studierenden zugeordnet. Die noch zu überprüfenden 25% sind nur schwer zu vermeiden, doch in 19 von 20 Fällen stimmt auch dort die vorgeschlagene Zuordnung und muss nur noch durch den Druck der Enter-Taste bestätigt werden.

Die Testreihen der Evaluationsphase ergaben einen zeitlichen Aufwand, der – konservativ gerechnet – bei 20% der Zeit für die Datenkontrolle mit der zuvor verwendeten proprietären Software liegt. Insgesamt bieten der TestBuilder und der TestReader im Vergleich also eine spürbare Erleichterung der Arbeitszeitbelastung und eine Reduzierung der notwendigen Bedienkompetenzen bei gleichbleibender Genauigkeit in der Erkennung der Antworten.

In der aktuellen Version liegt die Arbeitszeitbelastung einer Person für die eigentliche Auswertung somit bei ca. 4 Sekunden pro Klausur. 400 Klausuren sind damit in ca. 30 Minuten ausgewertet und überprüft, womit mehr Zeit für die Qualitätssicherung der Fragen bleibt.

Die Benutzeroberfläche der Programme ist dabei unkompliziert und intuitiv bedienbar.

Kontakt:

Dr. Thomas Lerche

Ludwig-Maximilians-Universität München

Lehrstuhl für Schulpädagogik

Leopoldstr. 13, 80802 München

thomas.lerche@edu.lmu.de

Ergänzender Hinweis:

Der TestReader enthält Software, die von der University of California, Berkeley und Beitragsleistenden entwickelt wurde.

Das Katzenicon wurde erstellt von Freepic (<http://www.freepik.com/free-photos-vectors/design>).
Veröffentlicht unter CC BY-ND