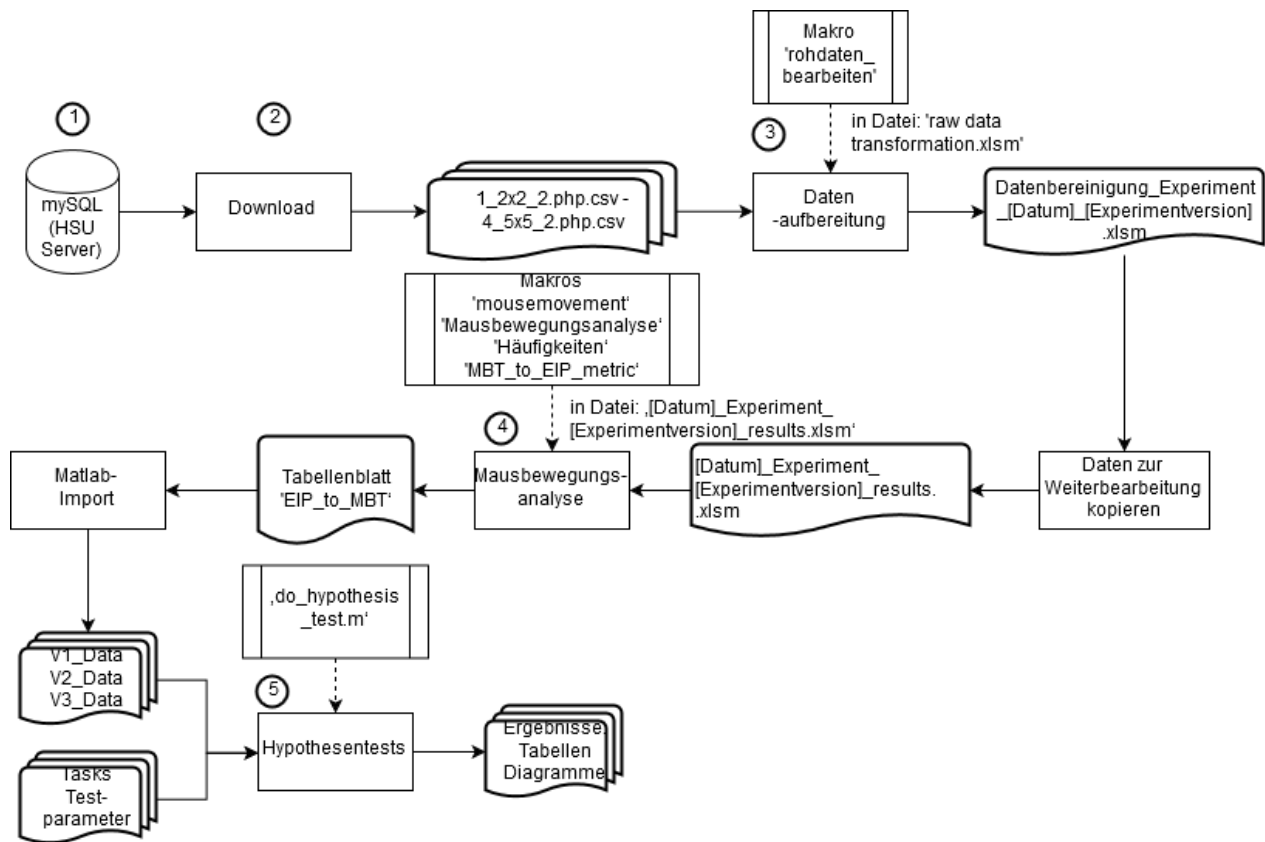


Vorgehen zur Datenanalyse

Eine Übersicht über das Vorgehen zur Datenanalyse ist in Abb. 1 gezeigt. Die Datenanalyse beginnt mit den Rohdaten aus dem Experiment, welche auf einem Universitätsserver in einer mySQL Datenbank gespeichert sind. In definierten Tätigkeiten (Rechtecke) werden, teilweise mithilfe von Tools (Makros im Falle der Bearbeitung in Excel und Skripte im Falle der Bearbeitung in Matlab; Prozesse), Zwischenergebnisse (Dokumente) erarbeitet. Das Vorgehen endet mit den graphisch aufbereiteten Ergebnissen der Hypothesentests (Tabellen und Diagramme).

Abb. 1 Übersicht zum Vorgehen bei der Datenanalyse.



Nachfolgend sind die fünf wesentlichen Schritte (markiert in Abb. 1) näher beschrieben:

1. Rohdaten sind in einer mySQL Datenbank für alle Spieler, gesammelt nach Experimentversion und Aufgabe, abgelegt.

2. Aus der mySQL Datenbank wird pro Kombination aus Experimentversion und Aufgabe ein Auszug im Format *.csv erstellt und zur weiteren Bearbeitung gespeichert (siehe Abb. 2). Jede csv-Datei enthält in der ersten Zeile einen Kopfstring, der folgende Spaltenüberschriften enthält: id,"expversion","transform","expname","subject","ip","choice","submitted","event","name","value","time","condnum". Jede weitere nichtleere Zeile enthält einen String, der aus der Aneinanderreihung der erfassten Verhaltensattribute gemäß der Spaltenüberschriften besteht.

Abb. 2 Ausschnitt aus CSV Datei '1_2x2_2.php.csv' angezeigt in Excel.

	A	B	C	D	E	F	G	H	I	J	K	L
1	id,"expversion","transform","expname","subject","ip","choice","submitted","event","name","value","time","condnum"											
2	1,"v2","","1_2x2_2.php","mattigol","95.91.242.245","btn2","2014-10-06 12:16:26","onload","body","body","990","-1"											
3	1,"v2","","1_2x2_2.php","mattigol","95.91.242.245","btn2","2014-10-06 12:16:26","subject","random","0","1014","-1"											
4	1,"v2","","1_2x2_2.php","mattigol","95.91.242.245","btn2","2014-10-06 12:16:26","order","col","0_1","1014","-1"											
5	1,"v2","","1_2x2_2.php","mattigol","95.91.242.245","btn2","2014-10-06 12:16:26","order","row","0_1","1015","-1"											

3. Die *.csv - Dateien werden über ein Makro in Excel von Dubletten und leeren Einträgen gereinigt und in ein Tabellenformat überführt (siehe Abb. 3). Die Ergebnistabellen zu allen Aufgaben einer Experimentversion werden in einer Datei ,Datenbereinigung_Experiment_[Datum]_[Experimentversion].xlsx' zusammengefasst, wobei jede Tabelle in einem eigenen Tabellenblatt abgelegt ist.

Abb. 3 Ausschnitt aus Überführungsdatei 'Datenbereinigung_Experiment 141006_V1.xlsx' angezeigt in Excel.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	id	expversion	transform	expname	subject	ip alt	ip	choice	submitted	event	name	value	time	
4115	658	v1	3_2x2_2.php	Derbolt		79239725	79.239.7.25		07.10.2014 14:17	onload	body	body	108	
4116	658	v1	3_2x2_2.php	Derbolt		79239725	79.239.7.25		07.10.2014 14:17	subject	random	1	120	
4117	658	v1	3_2x2_2.php	Derbolt		79239725	79.239.7.25		07.10.2014 14:17	order	col	0_1	120	
4118	658	v1	3_2x2_2.php	Derbolt		79239725	79.239.7.25		07.10.2014 14:17	order	row	1_0	120	

4. Der so bereinigte und gespeicherte Zwischenstand der Daten wird in die Datei ,[Datum]_Experiment_[Experimentversion]_results.xlsx' zur weiteren Auswertung kopiert. Die Daten werden über Makros hinsichtlich der verwendeten EIPs (bestimmt über eine Mausbewegungsanalyse gem. der EMMA-EIP-Interpretationsmetrik) ausgewertet und der über die gewählte Strategie erreichte Ergebniswert (payoff) bestimmt. Die Berechnung der Zwischenschritte erfolgt über die Makros ,mousemovement' → ,Mausbewegungsanalyse' → ,Häufigkeiten' → ,MBT_to_EIP_metric'. Die Ergebnisse der Zwischenschritte sind in den Tabellenblättern ,Mousemove' → ,Mausbewegungsanalyse' → ,Häufigkeiten' → ,EIP_to_MBT' abgebildet. Ein Ausschnitt der Ergebnisübersicht für Experimentversion V1 ist in Abb. 4 zu sehen.

Abb. 4 Ausschnitt aus Analyse-Datei '160804_Experiment_V1_results.xlsm' angezeigt in Excel.

	BP	BQ	BR	BS	BT	BU	BV	BW	BX	BY	BZ	CA	CB	CC	CD	CE
1	READ	FOCUS	CHOOSE	CHOOSE	END	COMPARE	COMPARE_I	ADD_I	ADD_II	STORE	EIP_Stream	P_Stream	sum_EIPs	sum_soph	time_EP	
2			1			1					MOVE_MOV b1MOVE_b1MOVE_a1MOVE_a1CHOOSE(I)_a1EN				0	
3						1					MOVE_END_a1MOVE_a1END				0	
4	9	9	1	9	1	1	3				5 MOVE_MOV a1MOVE_a0MOVE_a0OPEN_a0READ_a				2289	
5	5	5		5	1		2,5		1,5		4 MOVE_OPEN b1MOVE_a0OPEN_a0ELIMINATE_a0ADI				1348	
6	15	15	1	15	1	3	3		1		7 MOVE_OPEN a1MOVE_a1OPEN_a1READ_a1CHOOSE(I)				7390	
7	6	5	1	5	1	1	3				4 MOVE_MOV b1MOVE_b0MOVE_a0OPEN_a0READ_a				1056	
8	4	4	2	4	3	1	2,5		0,5		4 MOVE_MOV b1MOVE_a0MOVE_b0OPEN_b0READ_b				782	
9	4	4	1	4	2		1				1 MOVE_MOV b1MOVE_a0MOVE_b0OPEN_b0READ_b				752	
10			1		2		2				2 MOVE_MOV b1MOVE_a1MOVE_b1MOVE_b0COMP				0	
11	14	14	1	14	1	5	7			2	14 MOVE_OPEN a1MOVE_b1OPEN_b1READ_b1CHOOSE				1628	
12	5	5	1	5	1	1	3,5		1,5		6 MOVE_MOV c2MOVE_a0MOVE_b1MOVE_c1OPEN_c				251	
13	4	4	1	4	1	1	3,5		0,5		5 MOVE_MOV c2MOVE_b1MOVE_c1OPEN_c1READ_c1				643	
14	4	4	1	4	1	1	2,5		0,5		4 MOVE_MOV c2MOVE_b1MOVE_b1MOVE_b0OPEN_I				119	
15	17	17	1	17	1	4	4,5		1,5		10 MOVE_OPEN a3MOVE_a0OPEN_a0READ_a0CHOOSE(I)				2305	
16	3	3	1	3	1		3			1	4 MOVE_MOV b1MOVE_a1MOVE_b0OPEN_b0READ_b				1818	
17	4	4	1	4	1		1,5	1	0,5		3 MOVE_OPEN b1MOVE_a1OPEN_a0OPEN_b0OPEN_A				207	
18	9	7	1	7	2	3	0,5		0,5		4 MOVE_OPEN a1MOVE_a1OPEN_a1READ_a1ELIMINATE				618	
19	3	3	1	3	1	2	4				6 MOVE_OPEN a1MOVE_b1OPEN_b0OPEN_a0OPEN_B				1330	
20	5	3	1	3	1	1	2				3 MOVE_OPEN a2MOVE_a2OPEN_a2ELIMINATE_A2OP				579	
21	8	7	1	7	1	1	2				3 MOVE_OPEN c2MOVE_a0OPEN_a0ELIMINATE_b0OPI				1237	
22	8	8	1	8	1		1,5		0,5		2 MOVE_OPEN a2MOVE_a0OPEN_a0ELIMINATE_b0OPI				1135	
23					1						MOVE_END_b1MOVE_a1END				0	
24	7	7	1	7	1	1	1		1		3 MOVE_OPEN b1MOVE_a0OPEN_a0ELIMINATE_b0OP				958	

5. Das Ergebnis in Tabellenblatt ‚EIP_to_MBT‘ wird zur statistischen Auswertung als Tabelle in Matlab importiert, wo die Hypothesentests erfolgen. Das Modul ‚do_hypothesis_test.m‘ führt diese automatisiert für alle Experimentdaten (auch für alle Versionen) durch. Dabei wurde bei jedem Hypothesentest wie folgt vorgegangen:

1. Lade V[x]_data (← Tabelle enthält die gesamte Ergebnisübersicht aus der importierten xlsm-Datei für Experimentversion [x: 1,2,3]).
2. Wähle den für die jeweilige Hypothese relevanten Referenzdatensatz (entspricht Verhalten ohne Zeitdruck) und den Sample-Datensatz (vergleiche Kapitel 9 und Anhang E der Dissertation).
3. Führe pro Aufgabe und Zeitdruck-Parameterkonstellation t-Test bzw. portion test durch (~Hypothese; vergleiche Anhang E) und speichere Ergebnis.
4. Bereite Ergebnisse grafisch auf (Beispiele sind in Abb. 5 und Abb. 6 gezeigt).

Abb. 5 Darstellung der Populationsanteile gem. Hypothese IX und ihre Veränderung über die untersuchten Zeitdruck-Parameter (Ausschnitt aus Dissertation, S. 214).

FIGURE 43 – HYPOTHESIS IX: INFLUENCE OF TIME LIMITATION ON PROPORTIONS OF *RANDOM CHOICE*

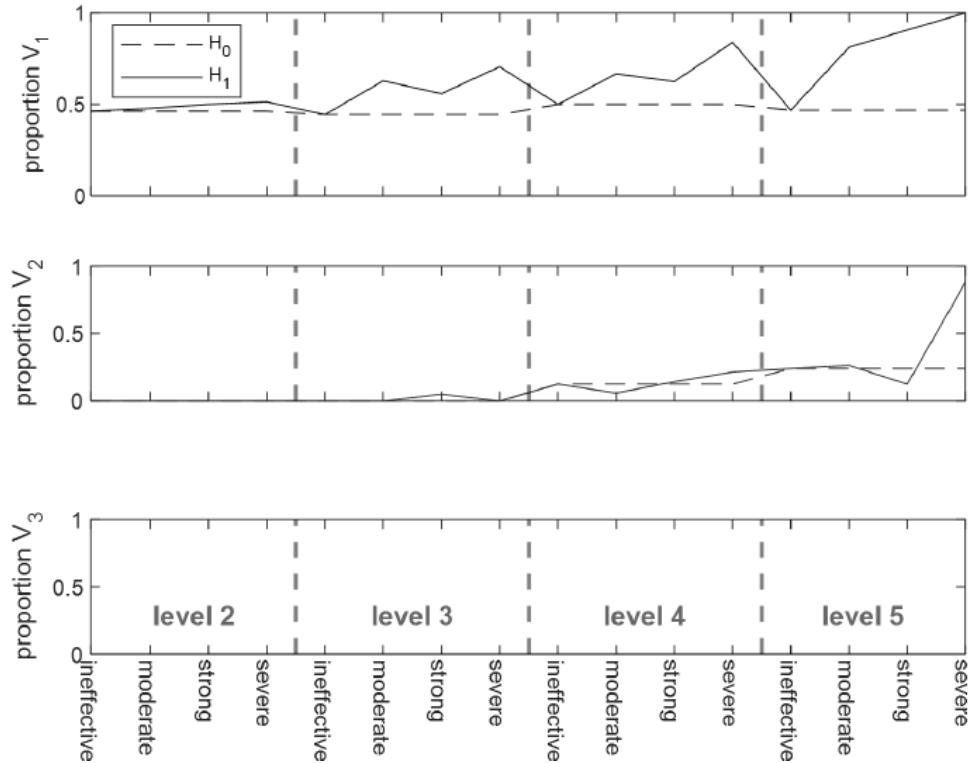


Abb. 6 Tabellarische Darstellung der ermittelten p-Werte zu Hypothese IX für die untersuchten Zeitdruck-Parameter (Ausschnitt aus Dissertation, S. 312).

TABLE 75 – *p*-VALUES OF HYPOTHESIS IX: INFLUENCE OF TIME LIMITATION

Level of complexity	Time limit [s]		<i>p</i> -value		
	V_1/V_3	V_2	V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
2	13	19	0.44	n.a. ¹⁸⁷	
2	7	9	0.36		
2	6	8	0.32		
3	75	75	H_0	H_0	H_0
3	35	56	0.06		