

Endbericht zum Projekt

UNTERSCHIEDUNG VON BESTANDS- UND FLUSSGRÖSSEN

Gefördert von:

Forschungskommission der Universität Klagenfurt
Kärntner Universitätsbund

Projektleitung:

Ao Univ. Prof. Dr. Günther Ossimitz

Mitarbeit am Projektbericht:

Ulrike Kotzent-Pietschnig, Benjamin Kreisler
Martin Waiguny, Melanie Zoltan

Interviews:

Sandra Galli, Eva Gratzner

Web-Umsetzung:

Renate Weichlinger

Oktober 2001
Universität Klagenfurt
Institut für Mathematik

<http://go.just.to/bf>

Inhaltsverzeichnis

1.	Ausgangslage _____	1-1
2.	Projektdesign _____	2-1
	2.1. Ziele des Projektes _____	2-1
	2.2. Erhebungsdesign _____	2-1
	2.3. Ergänzende Interviews _____	2-2
	2.4. Statistische Auswertungen _____	2-2
3.	Untersuchte Personen _____	3-1
4.	Ergebnisse des Pretests _____	4-1
	4.1. Ergebnisse Aufgabe A _____	4-1
	4.2. Ergebnisse Aufgabe B _____	4-3
	4.3. Ergebnisse Aufgabe C _____	4-5
	4.4. Ergebnisse Aufgabe D _____	4-7
	4.5. Ergebnisse Aufgabe E _____	4-10
	4.6. Zusammenfassung _____	4-11
5.	Ergebnisse des Posttests _____	5-1
	5.1. Ergebnisse Aufgabe F _____	5-1
	5.2. Ergebnisse Aufgabe G _____	5-2
	5.3. Ergebnisse Aufgabe H _____	5-4
	5.4. Ergebnisse Aufgabe K _____	5-6
	5.5. Ergebnisse Aufgabe L _____	5-8
	5.6. Zusammenfassung _____	5-10
6.	Vergleich Pretest – Posttest _____	6-1
	6.1. Zu- und Abfluss zeichnen _____	6-2
	6.2. Werte aus Tabelle lesen – Werte aus Graphik ermitteln _____	6-3
	6.2.1. Tabellendarstellung _____	6-3
	6.2.2. Graphikdarstellung _____	6-4
	6.2.3. Vergleich der Stories _____	6-4
	6.3. Bestand von gegebener Graphik zeichnen _____	6-8
	6.4. Bestand von verbaler Angabe zeichnen _____	6-9
	6.5. Zusammenfassung _____	6-10
7.	Ergänzende Interviews _____	7-1
	7.1. Ablauf _____	7-1
	7.1.1. Auswahl der befragten Personen _____	7-1
	7.1.2. Durchführung der Interviews _____	7-1
	7.1.3. Dokumentation der Interviews _____	7-1
	7.2. Interviews zum Vortest _____	7-2
	7.2.1. Aufgabe A _____	7-2
	7.2.2. Aufgabe B _____	7-4
	7.2.3. Aufgabe C _____	7-6
	7.2.4. Aufgabe D _____	7-7
	7.2.5. Aufgabe E _____	7-9
	7.2.6. Zusammenfassung der Interviews zum Vortest _____	7-11
	7.3. Interviews zu den Nachtests _____	7-11
	7.3.1. Aufgabe K _____	7-12
	7.3.2. Aufgabe L _____	7-13
	7.3.3. Aufgabe F _____	7-15
	7.3.4. Aufgabe G _____	7-15
	7.3.5. Aufgabe H _____	7-16
	7.3.6. Nachwort _____	7-18

8. Resümee	8-1
8.1. Inwieweit wurden die Ziele des Projektes erreicht?	8-1
8.2. Interessante Ergebnisse	8-3
9. Ausblick	9-1
10. Nachwort des Projektleiters	10-1
Literaturverzeichnis	1

Danksagung

Am Zustandekommen und an der erfolgreichen Durchführung des Projekts „Unterscheidung von Bestands- und Flussgrößen“ waren eine ganze Reihe von Personen maßgeblich beteiligt:

- Professor John Sterman vom Massachusetts Institute of Technology (MIT) hat mich mit einem bahnbrechendem Hauptvortrag an der System Dynamics Tagung 2000 in Bergen (Norwegen) auf das Phänomen aufmerksam gemacht, dass die Unterscheidung von Bestands- und Flussgrößen keine so einfache Sache ist.
- Die Forschungskommission der Universität Klagenfurt sowie der Kärntner Universitätsbund haben durch ihre finanzielle Unterstützung maßgeblich die erfolgreiche Durchführung dieses Projekts ermöglicht. Mein herzlichster Dank an die Verantwortlichen Prof. Helmut Rumppler und Dir. Hans Kampfer!
- Bei der Durchführung des Projekts haben eine Reihe von Studierenden der ABWL der Universität Klagenfurt engagiert mitgearbeitet. Im Rahmen einer Projektlehrveranstaltung haben Josef Biedermann, Ingrid Dogan, Edzo Frerichs, Andrea Grabner, Peter Ladinig, Reiner Micheler, Christian Petautschnig, Michael Rachbauer, Markus Steinbrugger und Philipp Tillian ganz entscheidend bei der Auswertung der einzelnen Teilfragen des Projekts weitestgehend ehrenamtlich mitgeholfen.
- Ganz besonderer Dank gebührt jedoch dem engeren Projektteam. Frau Ulrike Kotzent-Pietschnig und Herr Benjamin Kreisler haben als Stabsstelle hervorragende Koordinierungsarbeit geleistet und auch bei der Erstellung des Projektberichtes maßgeblich mitgewirkt und auch noch die Auswertung der Befragung unterstützt. Martin Waiguny und Melanie Zoltan engagierten sich hervorragend bei der Auswertung und beim Verfassen der Kernkapitel des Projektberichts, der auch von Herrn Waiguny gelayoutet wurde.
- Frau Sandra Galli und Eva Gratzler führten die ergänzenden Interviews mit viel Witz und Engagement durch, werteten diese aus und steuerten das entsprechende Kapitel im Projektbericht bei.
- Last but not least hat Frau Renate Weichlinger eine ansprechende Web-Site zum Projekt erstellt, die unter der Kurzadresse <http://go.just.to/bf> erreichbar ist.

Allen angeführten Personen gebührt mein herzlichster Dank. Ohne ihr Engagement wäre dieses Projekt nicht zustande gekommen.

1. Ausgangslage

(Günther Ossimitz)

Die Fähigkeit zu Systemischem Denken und Handeln wird in vielen Bereichen des wirtschaftlichen, öffentlichen und wissenschaftlichen Lebens als immer wichtiger angesehen¹. Dies kommt zunächst in einer umfangreichen Literatur zum Thema zum Ausdruck, die in verschiedensten Wissensgebieten (Wirtschaft/Management, Biologie, Systemdynamik, Kybernetik, Sozialwissenschaften, Psychologie, allgemeine Systemtheorie usw.) angesiedelt sind. Im Bereich Wirtschaft/Management ist es im deutschsprachigen Raum vor allem die St. Galler Managementsschule (Hans Ulrich, Peter Gomez, Gilbert J. Probst, auch Fredmund Malik), die seit Jahrzehnten konsequent systemorientierte Ansätze im Management vertreten (vgl. etwa Probst/Gomez 1991 oder Gomez/Probst 1987, 1997). International haben wohl die Überlegungen von Peter Senge (1990) zum Zusammenhang zwischen systemischem Denken und lernenden Organisationen die größte Verbreitung gefunden.

In der Psychologie ist es vor allem die Forschungsrichtung "Komplexes Problemlösen" mit den führenden Persönlichkeiten Dietrich Dörner und Joachim Funke, die in der deutschsprachigen Kognitionspsychologie umfangreiche empirische Untersuchungen über die Fähigkeit von Menschen zur Bewältigung komplexer Situationen anstellte. Die Untersuchungen wurden in der Regel an am Computer simulierten Szenarien durchgeführt. Eine gute Zusammenfassung der Ergebnisse bietet der Band "Die Logik des Misslingens" (Dörner 1989).

Im Bereich der systemdynamischen Modellierung mit moderner, grafisch orientierter Modellbildungswerkzeuge gibt es seit ca. 1987 einen Trend, die Fähigkeit zur systemischen Modellierung unter dem Label "systems thinking" zu vermarkten. Vorreiter dieser Linie ist Barry Richmond, seines Zeichens Mitentwickler der Systemdynamiksoftware Stella und Marketingdirektor der Stella-Vertriebsfirma HPS (vgl. etwa Richmond 1991, 1993, 1994).

¹ Für eine ausführliche Einführung in das Verständnis systemischen Denkens in verschiedenen Bereichen siehe Ossimitz (2000, S. 9-62).

Im Bereich Biologie-Umwelt ist es für den deutschen Sprachraum das Verdienst von Frederic Vester, durch öffentlichkeitswirksames Auftreten (mit Büchern, Fernsehdokumentationen, Computerspielen oder etwa der Wanderausstellung "Unsere Welt – ein vernetztes System") eine gewisse Sensibilisierung für das Thema "Systemisches Denken" in der Öffentlichkeit erreicht zu haben (vgl. etwa Vester (1983, 1988, 1999). Bei genauerer Durchsicht fällt jedoch auf, dass in dieser recht umfangreichen Literatur zwar recht viel von systemischen Denken, Systemdenken, vernetztem Denken, komplexem Problemlösen und ähnlichem die Rede ist, dass diese Begriffe jedoch kaum irgendwo griffig definiert worden sind. Am konkretesten ist noch Richmond mit seinen Erläuterungen zu "systems thinking", die jedoch inhaltlich enttäuschend bleiben: "systems thinking" sind für Richmond im wesentlichen diejenigen Fähigkeiten, die man braucht, um systemdynamische Simulationsmodelle zu entwickeln. Er sagt selbst: "Systems Thinking is System Dynamics with an aura" (Richmond 1994, S. 4).

Eine ausführlich dokumentierte und begründete Definition systemischen Denkens findet sich in Ossimitz (2000): "Entwicklung systemischen Denkens". Ossimitz unterscheidet vier mit einander zusammenhängende Dimensionen systemischen Denkens:

1. Denken in vernetzten Strukturen ("vernetztes Denken")
2. Denken in Zeitabläufen ("dynamisches Denken")
3. Denken in (bewusst wahrgenommenen) Modellen ("Modellorientiertes Denken")
4. praktisches Steuern von Systemen.

Jede dieser Dimensionen spielt eine Rolle, wenn systemisches Denken in einem umfassenden Sinn gefördert bzw. entwickelt werden soll. Allerdings werden bestimmte Maßnahmen zur Entwicklung systemischen Denkens nicht unbedingt alle Dimensionen gleichmäßig betreffen, sondern eher eine der genannten Dimensionen schwerpunktmäßig berücksichtigen.

Linda Booth Sweeney und John D. Sterman (Sweeney/Sterman 2000) haben in einer bahnbrechenden Studie "Bathtub Dynamics: Initial Results of an Systems Thinking Inventory" nachgewiesen, dass selbst hochqualifizierte und graduierte Studierende am renommierten Massachusetts Institute of Technology (MIT) erstaunliche Probleme

bei der Unterscheidung von Bestands- und Flussgrößen haben, wie sie schon für das Verständnis ganz elementare Vorgänge wie das Befüllen und Leeren einer Badewanne oder zum Verstehen des Unterschiedes zwischen Kontoständen und Kontobewegungen notwendig sind. Die Studie von Sweeney und Sterman zeigt, dass die Fähigkeit zur Unterscheidung von Beständen und Flüssen ("stock-flow-thinking") eine ganz zentrale, aber in der Praxis problematische Komponente dynamischen Denkens ist.

Die Untersuchungen von Sweeney und Sterman wurden von G. Ossimitz im Herbst 2000 mit einer Befragung von insgesamt 154 Studierenden der Universität Klagenfurt, der Universität Graz und der WU Wien (mit hauptsächlich wirtschaftsorientierten, z.T. auch naturwissenschaftlichen Studienrichtungen) fortgesetzt. Die Ergebnisse dieser Pilotstudie waren noch deutlich schlechter als in der Untersuchung von Sweeney/Sterman. Die wesentlichen Ergebnisse sind in Ossimitz (2001a, 2001b) zusammengefasst. Besonders bedenklich war das schlechte Abschneiden der Versuchspersonen bei Fragen, die auf die Unterscheidung zwischen Staatsverschuldung (eine Bestandsgröße) und Netto-Budgetdefizit (die zugehörige Flussgröße) abzielten. So meinten über 50% der Befragten, dass ein sinkendes Budgetdefizit eine sinkende Staatsverschuldung bedeute. (Tatsächlich steigert jedes noch so kleine Defizit die Staatsverschuldung und nur Budgetüberschüsse reduzieren die Staatsverschuldung).

Insgesamt zeigen die bereits durchgeführten Untersuchungen, dass häufig Flussgrößen für Bestandsgrößen gehalten werden und damit oft auch sehr einfache (tabellarisch oder graphisch gegebene) Daten falsch interpretiert werden. Dieses Defizit soll durch das vorliegende Projekt nochmals deutlicher herausgearbeitet und Möglichkeiten zu seiner Bekämpfung entwickelt werden.

2. Projektdesign

(Günther Ossimitz)

2.1. Ziele des Projektes

Mit dem Projekt "Unterscheidung von Bestands- und Flussgrößen" wurden folgende Ziele verfolgt:

- 1) Es sollte ermittelt werden, inwieweit Studierende der Betriebswirtschaft der Universität Klagenfurt in einem Pretest Aufgaben zur Unterscheidung von Bestands- und Flussgrößen bewältigen.
- 2) Nach dem Pretest sollte in einem "Crash-Kurs" von ca. 1–2 Stunden Dauer praktisch erprobt werden, inwieweit Basiskenntnisse zur Unterscheidung von Bestands- und Flussgrößen praktisch vermittelt werden können.
- 3) In einem Nachtest sollte festgestellt werden, inwieweit die Fähigkeit zur Unterscheidung von Bestands- und Flussgrößen sich durch den Crash-Test verbessert hat.

2.2. Erhebungsdesign

Aus diesen Zielen ergibt sich unmittelbar ein dreiphasiges Erhebungsdesign:

- Pretest (im März 2001)
- Crash-Kurs (Anfang Mai 2001)
- Posttest (Juni 2001)

Sowohl Pretest als auch Posttest erfolgten anonym durch einen schriftlichen Fragebogen. Die Versuchspersonen gaben bei beiden Erhebungen denselben Codenamen an, so dass bei Wahrung der Anonymität die Leistungsentwicklung einzelner Personen nachvollzogen werden kann.

Alle untersuchten Personen waren Teilnehmer(innen) einer Lehrveranstaltung "Angewandte Statistik für Betriebswirte", die ein Pflichtfach im ersten Studienabschnitt des Studiums der Angewandten BWL an der Universität Klagenfurt ist. Die Tests wie auch

der Crash-Kurs fanden im Rahmen der Vorlesungsstunden statt. Eine finanzielle Entschädigung der Probanden war nicht möglich. Diese Rahmenbedingungen brachten einige Einschränkungen: Am Nachtest nahmen nur mehr ca. $\frac{3}{4}$ der Personen teil, die den Vortest absolviert hatten. Weiters waren nicht alle Personen, die am Nachtest teilnahmen, auch tatsächlich beim Crash-Kurs anwesend.

2.3. *Ergänzende Interviews*

Ergänzend wurden freiwillige Versuchspersonen zu Interviews über Ihre Antworten auf den Fragebogen eingeladen. Dies bedingte ein Aufgeben der Anonymität. Beim Vortest waren insgesamt 13, beim Nachtest dann noch neun Personen zu einem Interview bereit. Beim Nachtest wurden nur mehr solche Personen interviewt, die auch zum Vortest interviewt wurden. Diejenigen Probanden, die beide Interviews absolvierten, erhielten eine kleine Entschädigung.

2.4. *Statistische Auswertungen*

Ungeachtet der relativ geringen Anzahl an Probanden wurde versucht, für verschiedene Fragen statistische Zusammenhänge zwischen dem Antwortverhalten und Variablen wie etwa dem Geschlecht nachzuweisen. Technisch wurde dazu jeweils ein Chi-Quadrat-Unabhängigkeitstest verwendet, der jeweils die Hypothese überprüft, dass die beiden Variablen unabhängig von einander sind. Da dieser Test jegliche Art von Abweichungen von der Unabhängigkeit überprüft, sind bei $n = 94$ bzw. $n=64$ (beim Nachtest) nur sehr schwer signifikante Abweichungen zwischen zwei Verteilungen zu erreichen. Bei so kleinen Werten von n liefert der Chi-Quadrat-Unabhängigkeitstest auch dann noch kein signifikantes Resultat, wenn tatsächlich die beiden Variablen nicht unabhängig von einander sind. Doch zur sicheren Diagnose, dass diese Abhängigkeit besteht, reicht in vielen Fällen ein Stichprobenumfang von unter 100 nicht aus. Nur bei ganz extremen Abhängigkeiten darf man bei so kleinen Stichprobenumfängen auch ein signifikantes Resultat erwarten.

3. Untersuchte Personen

(Benjamin Kreisler)

Das Projekt "Unterscheidung von Bestands- und Flussgrößen" wurde im Rahmen von Lehrveranstaltungen zur "Angewandte Statistik für Betriebswirte" an der Universität Klagenfurt im Sommersemester 2001 durchgeführt. Bei den untersuchten Personen handelte es sich ausschließlich um Studenten der Betriebswirtschaft. Ein Viertel der Teilnehmer hatte einen ähnlichen Test im vorangegangenen Wintersemester absolviert.

Nahmen am Pretest zu Semesterbeginn insgesamt 94 Hochschüler teil, so sank diese Zahl beim Postest drei Monate später auf 64. Dieser Umstand war allerdings bereits im Vorhinein evident, da das Ausscheiden einer gewissen Zahl an Studierenden aus einem Kurs während des Semesters durchaus natürlich ist.

Im Vorspann des Pretests gaben die Teilnehmer Geburtsjahr, Geschlecht und ihre Mathematiknote im Maturazeugnis an. Die jüngste Versuchsperson wurde 1982 geboren, der älteste Student war Jahrgang 1969. Der Großteil verteilt sich aber auf die Jahrgänge 1981 bis 1979 (20-22 Jahre).

Das Verhältnis zwischen männlichen und weiblichen VPn ist äußerst ausgewogen (53% : 47%), wobei sich die Zahlen beim Nachtest genau umkehren - zugunsten der Damen.

Auch bei den Mathematik-Maturanoten ist das Bild ausgeglichen: etwa je ein Fünftel aller Vpn hatten im Reifezeugnis ein "Sehr Gut", "Gut", "Befriedigend" bzw. "Genügend" in Mathematik im Maturazeugnis. Das letzte Fünftel hatte keine Mathematik-Note im Reifeprüfungszeugnis. Der einzige erkennbare "Trend" ist ein Zusammenhang zwischen Alter und Mathematiknote: Die "jüngeren" Jahrgänge (1978-82) hatten bei ihrer Matura eher bessere Mathematiknoten als die älteren. Ob dies die Folge besserer Unterrichtsmethoden oder eines gestiegenen Verständnisses für die Mathematik ist, oder einfach die Anforderungen gesunken sind, sei dahingestellt. Zwischen Geschlecht und Mathematiknote konnte hingegen kein Zusammenhang gefunden werden.

Weiters wurde jede Person vor Beginn des Pretests gebeten, auf einer 5-teiligen Skala eine Selbsteinschätzung bezüglich des Lesens bzw. Verstehens von Graphiken zu geben. Allgemein gab man sich vorsichtig-optimistisch: Einerseits bewerteten etwa 60% der Studenten ihre Fähigkeiten als nur "durchschnittlich", andererseits glaubten doch nur 12%, dass sie Graphiken schlecht bis gar nicht erfassen könnten. Eine Aufteilung der Antworten nach Geschlecht legt die Vermutung nahe, dass die männlichen Teilnehmer eher dazu neigen, ihre eigenen "skills" höher einzuschätzen als es die weiblichen tun. Das Ergebnis eines entsprechenden Chi-Quadrat-Unabhängigkeits-Tests war hochsignifikant.

Die letzten beiden Fragen des Vorspanns bezogen sich schon auf das eigentliche Thema der Untersuchung. Wortwörtlich wollten wir wissen: "Haben Sie in der Statistik schon einmal etwas über Bestandsmassen versus Bewegungsmassen gehört? – Ja oder Nein". Falls ja, sollte der Unterschied noch kurz erklärt werden. Dass gerade einmal 5 Studenten "Ja" angekreuzt haben, deutet einerseits sicher auf ein mangelndes Bewusstsein der Studierenden für dieses Problem hin – und das schon vor dem eigentlichen Test. Andererseits haben 95% gar keine Antwort gegeben. Dies ist wohl ein Indiz für die Schwierigkeit einer solchen Fragestellung. Man darf wohl vermuten, dass doch mehr als 5 von 94 Studenten das Konzept von Beständen und Flüssen während ihres Studiums schon einmal begegnet ist, nur vielleicht eben nicht mit genau den Bezeichnungen "Bestands-" und "Bewegungsmassen". Man kann annehmen, dass sich viele Studenten durch diese präzise vorgegebenen Begriffe verunsichern ließen und lieber keine als eine falsche Antwort gaben.

4. Ergebnisse des Pretests

(Martin Waiguny)

Der Pretest wurde von 94 Personen (50 männlich, 44 weiblich) ausgefüllt. 23 der Teilnehmer hatten bereits im Semester davor einen ähnlich gestalteten Test ausgefüllt.

4.1. Ergebnisse Aufgabe A

- A) Eine Regentonne mit 100 l Fassungsvermögen ist um 12:00 Uhr leer. Um exakt 14:00 Uhr beginnt es leicht zu regnen und aus der Dachrinne füllt sich die Tonne mit 25l/Stunde Zufluss bis Mitternacht. Die Tonne besitzt keinen Abfluss, sondern wird so lange gefüllt, bis sie übergeht. Das übergehende Wasser wird als Abfluss gewertet. Zeichnen Sie den Zufluss und den Abfluss zwischen 12:00 Uhr und 00:00 Uhr in die Diagramme unten ein!

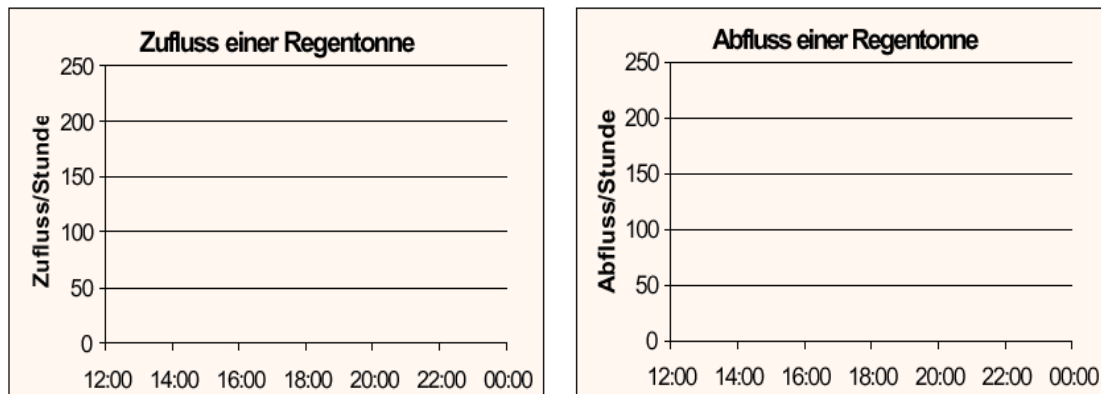


Abbildung 4-1: Aufgabe A - Angabe

Wie aus der Angabe in Abbildung 1 ersichtlich, mussten die Versuchspersonen (VPn) zunächst erkennen, dass von 14:00 bis 0:00 Uhr ein konstanter Zufluss von 25 Litern pro Stunde besteht, und diesen in das Diagramm einzeichnen.

Nur 34 der 94 Personen erkannten korrekt, dass es sich um eine konstant bleibende Zuflussgröße handelt. Eine Mehrheit von 52 VPn zeichnete den im Laufe der Zeit ab 14:00 linear steigenden Bestand an Wasser in der Regentonne ein. Einige Versuchspersonen berücksichtigten dabei nicht, dass die Tonne nur 100 l Fassungsvermögen hat und zeichneten den Bestand auch über die Grenze von 100 l (die nach vier Stunden Zufluss um 20:00 erreicht wird) hinaus ein. Man kann dies so interpretieren, dass diese Personen den kumulierten Zufluss im Laufe der Zeit in das Diagramm

eingetragen haben. Als wahrhaft kurios muss man die Lösungen von acht VPn bezeichnen, bei denen der Bestand in Form einer Treppenfunktion eingezeichnet wurde; so, als ob sich jede Stunde genau mit dem Gongschlag die Wassermenge schlagartig um 25 Liter erhöhen würde. Wo das Wasser dafür plötzlich herkommt, lässt sich leider nicht erahnen. Auch mit der Uhrzeit hatten einige Teilnehmer ihre Probleme und zeichneten den Zufluss (bzw. Bestand) bereits ab 12:00 Uhr ein.

Eine Besonderheit ist auch, dass alle 5 komplett richtigen Antworten nur von männlichen VPn stammen. Aufgrund der geringen absoluten Zahl ist dies jedoch noch kein signifikantes Ergebnis.

Im Diagramm "Abfluss einer Regentonne" sollten die VPn einen konstanten Abfluss von 25 Liter/Stunde zwischen 18:00 und 24:00 Uhr einzeichnen.

Hier waren die Teilnehmer etwas erfolgreicher als beim Erkennen des Zuflusses. Immerhin zeichneten elf Personen einen vollständig richtigen Graphen ein. Erschütternd ist allerdings, dass die Zahl der annähernd richtigen Graphen deutlich geringer war als bei der Zufluss-Teilaufgabe. Erkannten bei der ersten Teilaufgabe noch 34 Versuchspersonen, dass es sich um einen Zufluss und nicht um einen Bestand handelte, waren dies beim Abfluss nur mehr 12. Auch im zweiten Teil hatten die Teilnehmer ihre Probleme mit der Uhrzeit und zeichneten oft auch einen Abfluss von 100 Litern je Stunde ein.

Tabelle 4-1: Antworten Aufgabe 1, Teilaufgabe Abflüsse

Geschlecht	komplett richtig	Flüsse richtig, falscher Zeitpunkt	Zeit richtig, Abfluss falsch	Falsche Zeit, falscher Abfluss vom Ansatz her richtig	Richtige Zeit, sonst falsch	total falsch	Gesamt
m	6		24	3	6	11	50
w	5	1	12	2	6	18	44
Gesamt	11	1	36	5	12	29	94

Tabelle 1 zeigt die Verteilung der Antworten für die Teilaufgabe "Abfluss aus der Regentonne", gegliedert nach Geschlecht. Wie bei der Teilaufgabe "Zuflüsse" gibt es auch hier keinen signifikanten Zusammenhang zwischen Antwortverhalten und Geschlecht. Dafür hat die Anzahl der komplett falschen Antworten stark zugenom-

men (8 beim Zufluss, 29 beim Abfluss). Man könnte dadurch fast zum Schluß kommen, dass viele Versuchspersonen einfach nur irgend einen Graphen "nach Gefühl" eingezeichnet haben – zumindest lässt die kuriose Vielfalt an Treppenfunktionen, Geraden, Kurven und ähnlichem darauf schließen.

4.2. Ergebnisse Aufgabe B

- B) In der Ski-Saison werden viele Verletzte aus der Umgebung in das Krankenhaus Schwarzach-St.Veit eingeliefert. An den Hauptschitagen müssen die meisten neuen Patienten aufgenommen werden. Entlassungen finden hauptsächlich an Wochentagen statt. Durchschnittlich bleiben die Patienten 5-12 Tage in stationärer Behandlung. Die folgende Statistik zeigt für zwei Wochen für jeden Tag die Zahl der neu aufgenommenen und entlassenen Patienten:

	Datum	Aufnahmen	Entlassungen	Datum	Aufnahmen	Entlassungen
So	31.12.2000	13	0	07.01.2001	19	3
Mo	01.01.2001	18	0	08.01.2001	11	9
Di	02.01.2001	10	4	09.01.2001	12	15
Mi	03.01.2001	7	5	10.01.2001	8	17
Do	04.01.2001	10	7	11.01.2001	10	25
Fr	05.01.2001	10	8	12.01.2001	8	28
Sa	06.01.2001	12	0	13.01.2001	14	30

Alle folgenden Fragen beziehen sich auf den in der Tabelle erfassten Zeitraum!

- B1 An welchem Tag waren die meisten Patienten im Spital?
Tag: _____ bzw. lässt sich nicht beantworten, weil....
- B2 Wie sind Sie auf Ihre Antwort in Aufgabe B1 gekommen?
- B3 An welchem Tag haben die meisten Patienten das Spital verlassen?
Tag: _____ bzw. lässt sich nicht beantworten, weil....
- B4 An welchem Tag sind die meisten Betten frei geworden?
Tag: _____ bzw. lässt sich nicht beantworten, weil....

Abbildung 4-2: Aufgabe B - Angabe

Hier sollte aus der Tabelle erkannt werden, dass der Bestand an Patienten in der Nacht vom 8.1. auf den 9.1. am höchsten ist². Dies kann man durch folgende Überlegung auch ohne genaues Nachrechnen erkennen: Bis einschließlich 8.1. wurden an jedem Tag mehr Patienten aufgenommen als entlassen, daher musste die Zahl der Patienten am Ende jedes Tages höher sein als zu Tagesbeginn. Ab 9.1. ist es genau umgekehrt: jeden Tag werden mehr Patienten entlassen als neu aufgenommen worden sind. Damit sinkt die Zahl der Patienten laufend. Insgesamt

² In der Auswertung wurde jede der Antworten 8.1. oder 9.1 sowie auch 8.1./9.1. als richtig gewertet.

folgt daraus, dass in der Nacht von 8.1. auf 9.1. die meisten Patienten im Spital waren. Weiters musste die Antwort auf Teilfrage B1 sowie alle Antworten "das lässt sich nicht genau sagen" begründet werden. Außerdem musste der höchste Bruttoabfluss (13.1.) sowie der höchste Nettoabfluss (12.1.) ermittelt werden, was sich auch sehr einfach aus der Tabelle ablesen bzw. errechnen lässt.

57 (60,6 %) VPn errechneten durch Addieren und Subtrahieren der Aufnahmen und Entlassungen korrekt den höchsten Bestand an Patienten zum 8.1. Insgesamt könnte man sogar noch von einer höheren Anzahl an richtigen Antworten ausgehen, da weitere 10 TN sich durch Rechenfehler oder Fehleinschätzungen irrten. Immerhin 60 TN erkannten, dass der maximale Bestand mit der Gegenüberstellung von Aufnahmen und Entlassungen in Zusammenhang steht. Auch wenn die meisten der korrekten Antworten von männlichen VP kamen, so war auch hier die Abhängigkeit vom Geschlecht nicht signifikant. Die meisten der 15 VP, die angaben "Die Frage lässt sich nicht beantworten" begründeten dies mit dem Fehlen eines Anfangsbestandes, der ihrer Meinung nach nötig wäre, um den laufenden Bestand zu errechnen.

Weit weniger Probleme hatten die TN mit der Bestimmung der höchsten Zahl an Entlassungen (Frage B3). Hier kamen immerhin 89 TN (94,7 %) auf die korrekte Antwort.

Die Frage B4 "Wann wurden die meisten Betten frei" bereitete hingegen mehr VPn Schwierigkeiten. 34 VPn waren der Meinung, dass es sich um dieselbe Frage wie bei B3 handelte. Man kann natürlich argumentieren, dass für jeden Patienten, der das Spital verlässt, auch ein Bett frei wird und in diesem Sinne hätten Frage B3 und B4 tatsächlich dieselbe Antwort, nämlich, dass am 13.1. die meisten (30) Entlassungen stattfanden. Intendiert war allerdings mit Frage B4, wann per Saldo die meisten Betten frei wurden. Insgesamt wird man die Ergebnisse bei B4 wohl zurückhaltend interpretieren müssen, weil sich zwei unterschiedliche Antworten plausibel argumentieren lassen. Mehr kurios als plausibel waren hingegen einige der Begründungen von VPn, die meinten, Frage B4 ließe sich nicht beantworten, wie etwa: "Es fehlt die Angabe ob die Patienten stationär oder ambulant behandelt werden. "

4.3. Ergebnisse Aufgabe C

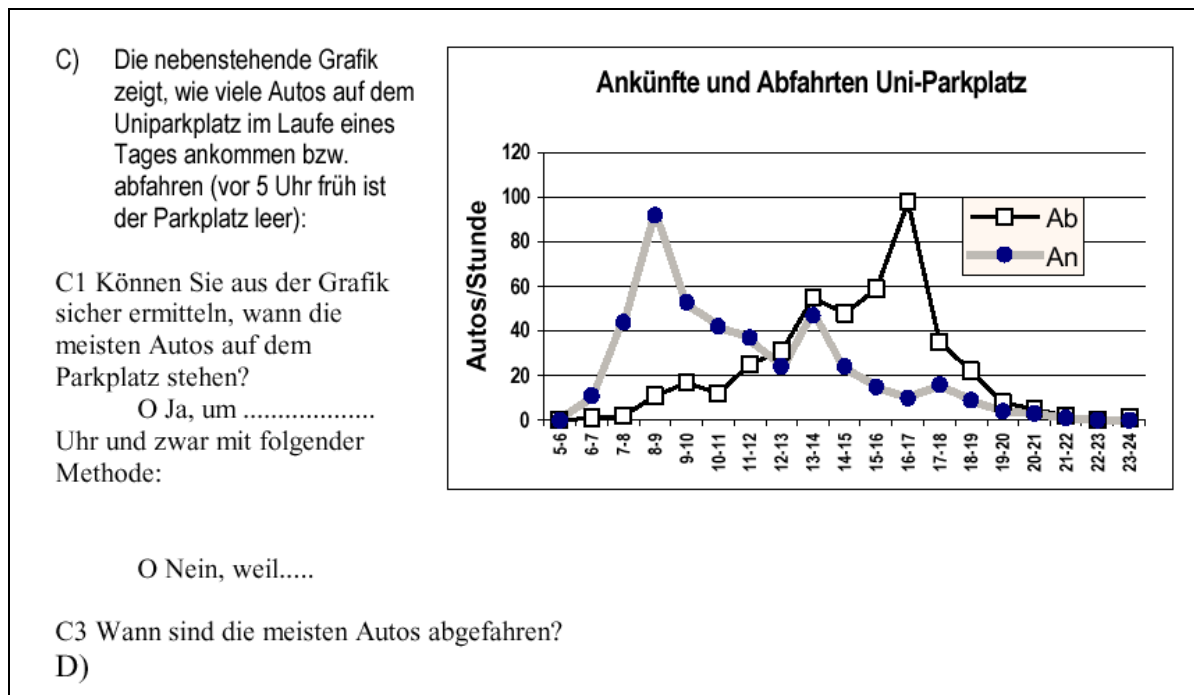


Abbildung 4-3: Aufgabe C - Angabe

Dieses Beispiel ist von der Aufgabenstellung der Aufgabe B sehr ähnlich. Wiederum sollte der maximale Bestand an Autos auf dem Uniparkplatz festgestellt werden, sowie die größte Abflussgröße. Der Unterschied zu Aufgabe B bestand in der Darstellung der Flussgrößen als Graphik. Die richtige Lösung für den maximalen Bestand liegt genau im Schnittpunkt der Kurven (etwa um 12 Uhr³). Die Argumentation läuft analog wie bei Aufgabe B: vor 12 Uhr kommen in jeder Stunde mehr Autos an als abfahren, daher muss die Zahl der Autos am Ende der Stunde größer sein als zu Beginn der Stunde. Umgekehrt fahren nach 12:00 Uhr jede Stunde mehr Autos ab als ankommen, und somit sinkt die Zahl der Autos von voller Stunde zu voller Stunde. Zusätzlich sollten die Teilnehmer auch eine Begründung für ihre Lösung abliefern. Die meisten Abfahrten waren wiederum aus der Abfahrtenkurve ersichtlich und diese erreichte in der Stunde von 16:00 bis 17:00 Uhr ihren Hochpunkt.

Mit dem Erkennen des Bestandes hatten bei dieser Aufgabe wesentlich mehr Teilnehmer Probleme als bei der Tabellendarstellung, wo viele VPn noch durch simples Ausrechnen das korrekte Maximum bestimmen konnten. Die Graphik wurde

³ Wie bei Aufgabe B wurde wiederum jede Angabe zwischen 11 und 13 Uhr als richtig gewertet.

nur von 17 Personen richtig interpretiert, die im Schnittpunkt den maximalen Bestand erkannten. Der Großteil der Teilnehmer, 54 Versuchspersonen, nahmen den Hochpunkt der Ankunftscurve bzw. 2 Versuchspersonen nahmen den Hochpunkt der Abfahrtskurve als Bestandsmaximum an. Die meisten VPn hielten also den Zufluss irrtümlich für den Bestand.

Interessant sind auch die Begründungen für die Lösungen. Selbst bei den richtigen Lösungen wurde oft nicht die korrekte Begründung gegeben. Der Großteil der Teilnehmer nahm den größten Abstand zwischen den Kurven an. Eine häufige Antwort lautete somit "Ankünfte minus Abfahrten". Eigentlich ist diese Antwort von der Grundidee her richtig: Wenn der Nettofluss 0 beträgt; also Zufluss und Abfluss gleich groß sind (Schnittpunkt), ist der größte Bestand erreicht. Dennoch meinten mit dieser Antwort die meisten Teilnehmer dennoch den Abstand zwischen den Kurven. Einige gaben direkt an, dass die meisten Ankünfte bzw. Abfahrten für den maximalen Bestand ausschlaggebend wären.

Erstaunlich ist auch der Zusammenhang zwischen Antwort auf die Frage C1 und der Eigenbeurteilung bezüglich des Lesens von Graphiken. Insgesamt haben sich 27 Versuchspersonen als "sehr gut" oder "gut" eingeschätzt. Von diesen haben allerdings nur vier dieser Teilnehmer eine richtige Antwort abgegeben. Von den 54 VPn, die sich mit "durchschnittlich" bewerteten, kamen immerhin auch zehn Personen auf das richtige Ergebnis. Insgesamt zeigt sich, dass diejenigen, die sich im Lesen von Graphiken als "sehr gut" oder "gut" einschätzten, keineswegs Aufgabe C1 besser bewältigten als die übrigen.

Fünfzehn Teilnehmer gaben an, dass diese Aufgabe nicht lösbar sei. Viele begründeten dies damit, dass die Graphik zu ungenau skaliert wäre bzw. das sie zu kompliziert wäre oder dass die Graphik keine Bestände angebe.

Mit der zweiten Aufgabenstellung C3 hatten die Teilnehmer weit weniger Probleme. 83 Teilnehmer gaben hier die korrekte Lösung an. Hier kann man also einen Rückschluss ziehen, dass bei einem graphisch gegebenem Abfluss das Maximum dieses Abflusses relativ sicher erkannt werden kann. Zusätzlich untersuchten wir, ob für

Bei der Frage der maximalen Abflussgröße ein Zeitintervall oder ein Zeitpunkt angegeben wurde. Hier entschieden sich 81 für ein Zeitintervall und 12 für einen Zeitpunkt. Dies kann man auf die Angabe zurückzuführen, wo bei der Skalierung ebenfalls Intervalle verwendet werden.

4.4. Ergebnisse Aufgabe D

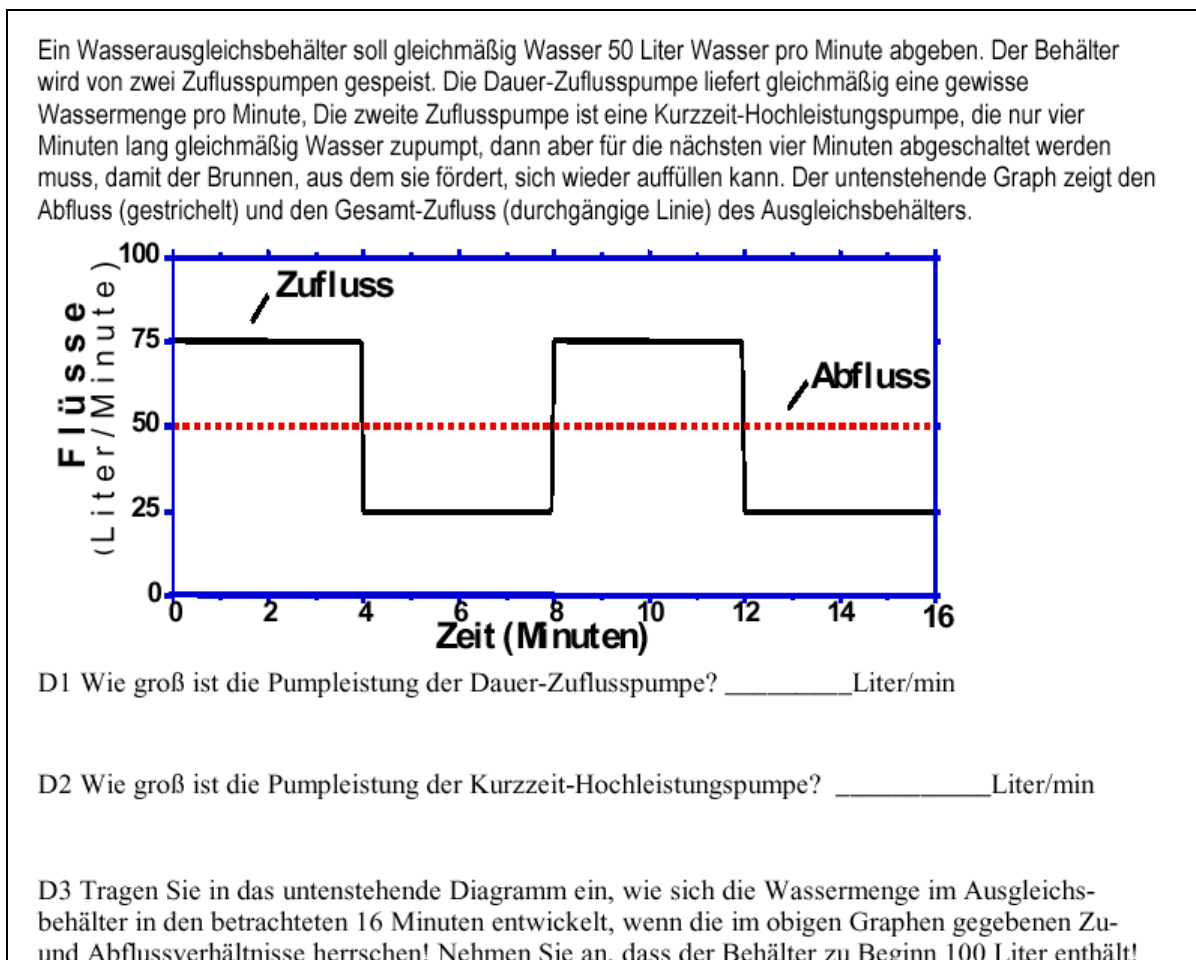


Abbildung 4-4: Aufgabe D - Angabe

Diese Aufgabe ist der klassischen Badewannen-Aufgabe aus "Bathtub Dynamics" von Sweeney/Sterman(2000) sehr ähnlich. Die VPn mussten hier zuerst die Leistungen (Zuflussgrößen) der Druckpumpen erkennen und anschließend in ein vorgefertigtes Diagramm den Bestand an Wasser im Wasserausgleichsbehälter einzeichnen. Die richtige Lösung ist aus Abbildung 4-4 ersichtlich.

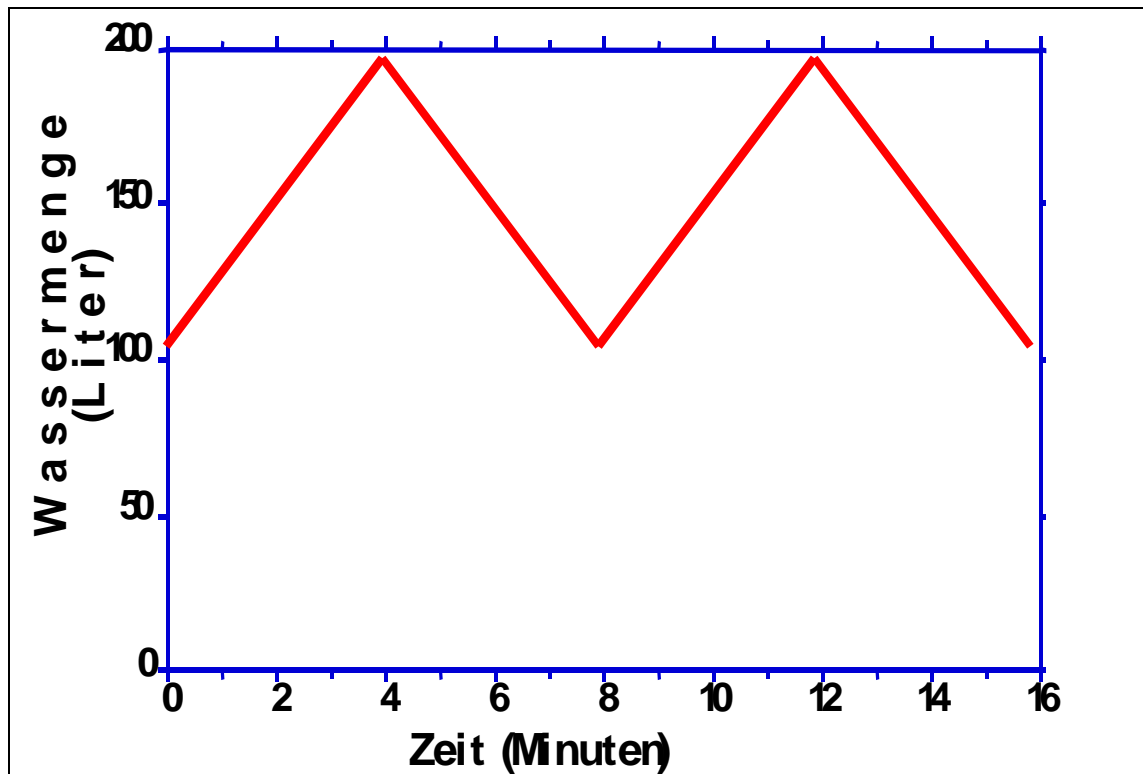


Abbildung 4-5: Musterlösung Aufgabe D

Die Leistung der Pumpen kann direkt aus der Graphik abgelesen werden. Die Leistung der Dauerpumpe beträgt konstant 25 Liter/Minute, wie 57 der VPn richtig erkannten. Die Leistung der Kurzzeit-Hochleistungspumpe ist einfach die Differenz auf den Gesamtzufluss von 75 Liter/Minute, also 50 Liter/min. Dies konnte ebenfalls aus der Graphik recht einfach ermittelt werden und wurde von 52 VPn korrekt erkannt.

Das Zeichnen des Bestandsgraphen, der eine Zickzackkurve mit den Hochpunkten bei 4 Minuten und bei 12 Minuten und den Tiefpunkten bei 8 Minuten und 16 Minuten darstellt fiel den meisten VPn recht schwer.

Es wurde auch noch untersucht, inwieweit die einzelnen Lösungen folgende Kriterien erfüllen:

- Ist der Zufluss größer als der Abfluss, steigt die Wassermenge
- Ist der Abfluss größer als der Zufluss, sinkt die Wassermenge
- Der Graph soll keine diskontinuierlichen Sprünge enthalten (er ist stetig)

- Die Extremwerte des Graphen treten dann auf, wenn der Nettofluss gleich null ist ($t=4, 8, 12, 16$)
- In jedem Anschnitt ist der Nettofluss konstant, der Bestand muss linear steigen oder fallen
- Die Maxima liegen bei 200, die Minima bei 100 Liter Bestand.

Waren alle diese Kriterien erfüllt so wurde die Graphik als richtig eingestuft. Nur 15 Lösungen waren vollkommen korrekt.

Aus dem Erfülltsein bzw. Nicht-Erfülltsein einzelner der angeführten Kriterien wurde versucht, auch die nicht vollständig korrekten Lösungen in Gruppen zu klassifizieren. Neben den 15 vollkommen korrekten Lösungen waren weitere 14 Lösungen (hinsichtlich des symmetrischen Zick-Zackmusters) beinahe korrekt. Vielen der VPn gelang es jedoch nicht, von den Flüssen auf die Wassermenge (den Bestand an Wasser) zu schließen. 27 Teilnehmer zeichneten als Bestand einen ähnlichen Graphiktyp wie in der Angabe mit vertikalen Sprüngen zu den Zeitpunkten 4, 8, 12. Fünf VPn entschieden sich überhaupt für horizontale Geraden, was der Idee entspricht, dass sich der Wasserstand sich überhaupt nicht ändert⁴. Die restlichen 33 Teilnehmer lieferten Graphen, die nur als total falsch klassifiziert werden konnten oder gaben gar keine Lösung an.

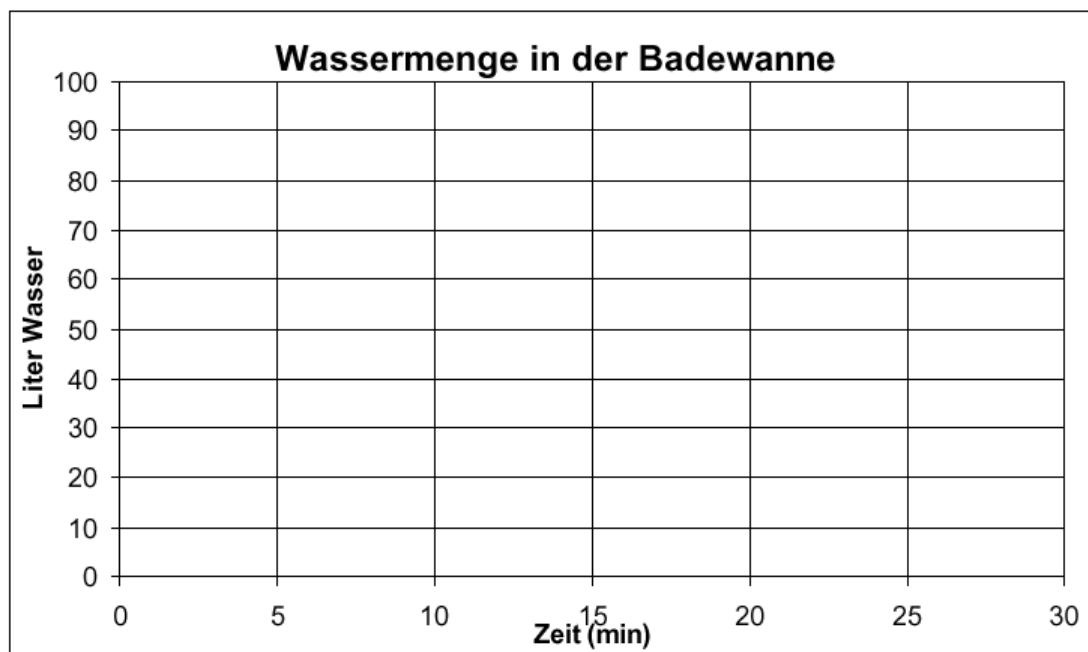
Erstaunlich ist, dass 4 von 10 VPn, die sich bei ihrem Graphikverständnis mit "schlecht" eingestuft haben, dennoch eine richtige Lösung abgaben. Von den Personen, die sich bezüglich ihrer Fähigkeit, Graphiken zu lesen, als "sehr gut" bzw. "gut" einschätzten, waren nur 25% bzw. 21% aller Lösungen korrekt.

⁴ Eine VPn gab mit einer solchen Lösung gab im anschließenden Interview an, dass sie den Bestand an Wasser als völlig unabhängig von den Zu- und Abflüssen sieht.

4.5. Ergebnisse Aufgabe E

E)

Herr Maier lässt um Punkt 19:00 Uhr Wasser in seine Badewanne einlaufen. Die Zuflussmenge beträgt konstant 14 Liter/min. Um exakt 19:04 Uhr bemerkt Herr Maier, dass der Abfluss der Wanne offen war und schließt den Abfluss. Aus dem Abfluss fließen pro Minute konstant 9 Liter Wasser. Um exakt 19:09 schließt Herr Maier den Wasserzufluss und genießt bis 19:15 die gefüllte Wanne. Um exakt 19:15 öffnet er den Abfluss und lässt das gesamte Wasser ausrinnen.



E1 Skizzieren Sie in obiger Grafik in etwa, wie sich die Wassermenge in der Badewanne im Laufe der Zeit entwickelt! Schwierigkeitsgrad (1 very easy 5 very heavy): _____

Abbildung 4-6: Aufgabe E - Angabe

Diese Frage war von der Aufgabenstellung her ähnlich wie die in Aufgabe D. Wiederum mussten die Versuchspersonen den Bestand an Wasser einer Badewanne einzeichnen, diesmal stammen die Daten allerdings aus einer verbalen Angabe. Der Bestand wächst von 19:00 Uhr bis 19:04 Uhr konstant per Saldo um 5 Liter pro Minute auf insgesamt 20 Liter um 19:04 Uhr. Ab 19:04 Uhr beträgt der Zufluss fünf Minuten lang 14 Liter/Minute (also insgesamt 70 Liter) bis zu einem Gesamtwasserstand von 90 Litern. Ab 19:15 verringert sich der Bestand dann konstant um 9 Liter pro Minute und die Wanne ist somit exakt zehn Minuten später leer. Hier

mussten die Versuchspersonen für die ersten 4 Minuten einen Nettozufluss ausrechnen und dann einen konstanten Zufluss und einen konstanten Abfluss bestimmen.

Erstaunlicherweise waren hier die VPn wesentlich erfolgreicher als bei Aufgabe D, obwohl die Aufgabenstellung sehr ähnlich ist. 51 Teilnehmer gaben eine völlig korrekte Lösung an und weitere acht hatten nur kleinere Rechenfehler. Aber auch weitere 19 Testpersonen hatten vom Ansatz her eine richtige Lösung, obwohl bereits gröbere Fehler eingebaut wurden. 16 VPn gaben auf Frage E entweder keine oder eine komplett falsche Lösung ab.

Aus dem Vergleich zu Aufgabe D darf man vorsichtig schließen, dass die VPn wesentlich weniger Probleme haben, eine verbale Angabe zu interpretieren, als graphisch gegebene Zu- und Abflüsse korrekt in Bestände umzuwandeln. Dennoch sind die beiden Aufgaben D und E nicht unabhängig von einander: Immerhin konnten 14 der 15 VPn, die bereits Aufgabe D richtig gelöst hatten, auch Beispiel E korrekt lösen.

4.6. Zusammenfassung

Abschließend wurde von uns untersucht, wie viele Aufgaben die einzelnen VPn im Pretest vollständig richtig gelöst haben. Diese Ergebnisse waren ernüchternd. 77 von 94 Vpn (82%) schafften kein einziges der fünf Aufgaben A-E vollständig richtig. Zwei Personen schafften ein Beispiel korrekt, immerhin 13 Personen (14%) lösten wenigstens zwei der fünf Aufgaben korrekt. Drei bzw. vier korrekt gelöste Beispiele schafften nur jeweils eine Person. Etwas freundlicher fällt das Ergebnis aus, wenn man die fünf Aufgaben in 13 Teilantworten zerlegt (Abbildung 4-6). Man erhält insgesamt eine recht schöne Normalverteilung mit einem Maximalwert bei 6 richtig gelösten Teilaufgaben. Die Anzahl der erfolgreich gelösten Teilaufgaben war mit keiner der Variablen Alter, Geschlecht, Maturanote und Selbsteinschätzung zum Graphikverständnis signifikant korreliert.

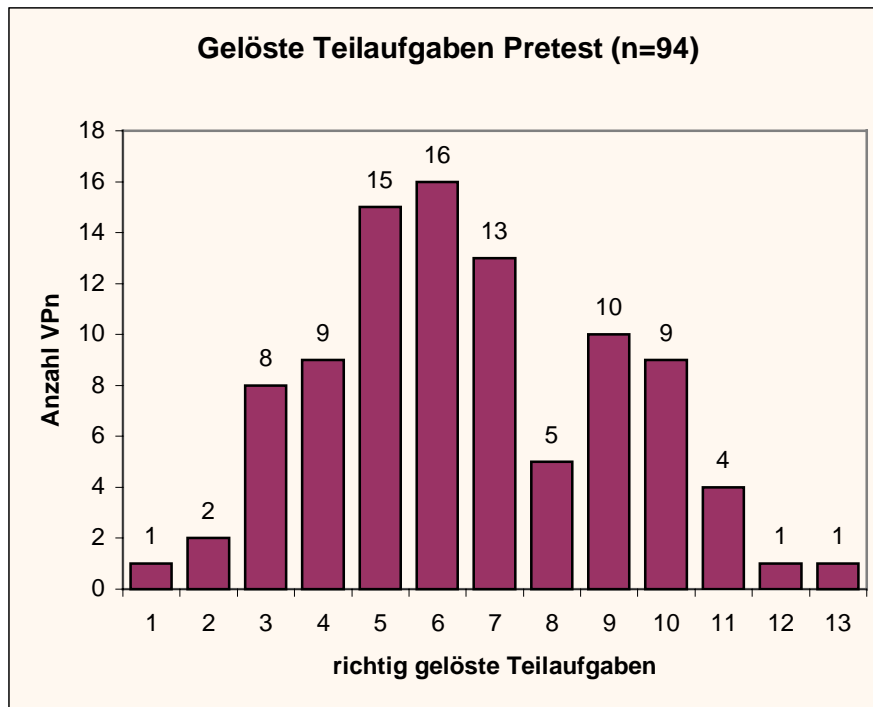


Abbildung 4-7: Häufigkeitsverteilung der Antworten

Abschließend sei noch erwähnt, dass man die Ergebnisse des Pretests auch unter der Tatsache betrachten muss, dass weder Sanktions- noch Honorierungsmöglichkeiten für die Teilnehmer vorgesehen waren. Viele mögen sich daher nicht richtig angestrengt haben und daher könnte es auch zu diesen doch recht ernüchternden Ergebnissen kommen, allerdings haben vorangegangene ähnliche Untersuchungen von Sweeney/Sterman (2000) und Ossimitz (2001a, 2001b) auch keine wesentlich anderen Ergebnisse gebracht.

Man könnte aufgrund all dieser Tests bemerken, dass eigentlich viele Leute mit komplexeren Graphikproblemstellungen Schwierigkeiten haben oder einfach zu schlampig bei den Lösungen und beim Ablesen der Daten sind. Verbale Angaben und auch tabellarische Angaben verursachen weit weniger Probleme wie die Ergebnisse doch recht eindeutig gezeigt haben.

5. Ergebnisse des Posttests

(Melanie Zoltan)

Der Posttest wurde etwa fünf Wochen nach dem Crash Kurs durchgeführt. Die Aufgaben waren denen beim Pretest sehr ähnlich. Dieser Test wurde nur mehr von 64 Personen (30 männlich, 34 weiblich) ausgefüllt.

5.1. Ergebnisse Aufgabe F

- F) Eine Regentonne mit 100 l Fassungsvermögen ist um 12:00 Uhr leer. Um exakt 14:00 Uhr beginnt es leicht zu regnen und aus der Dachrinne füllt sich die Tonne mit 25l/Stunde Zufluss bis Mitternacht. Die Tonne besitzt keinen Abfluss, sondern wird so lange gefüllt, bis sie übergeht. Das übergehende Wasser wird als Abfluss gewertet. Zeichnen Sie die Wassermenge in der Regentonne zwischen 12:00 Uhr und 24:00 Uhr ein!

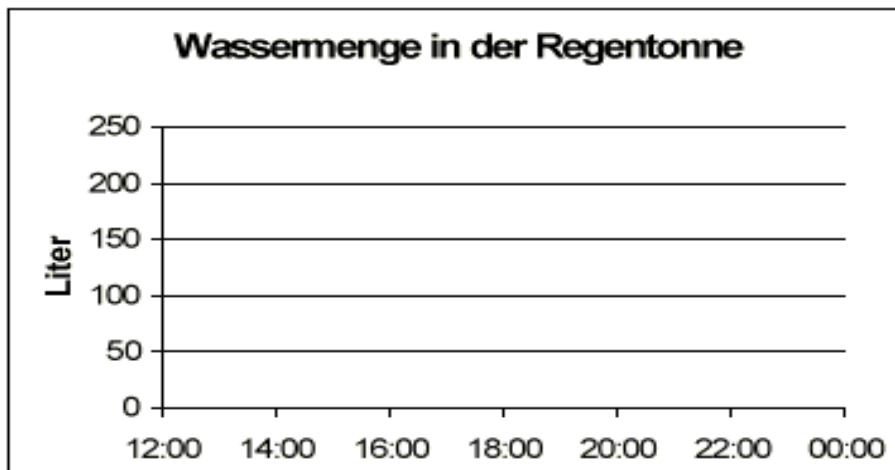


Abbildung 5-1: Aufgabe F - Angabe

Bei dieser Aufgabe geht es darum, die Wassermenge - also eine Bestandsgröße - einzuzeichnen. Der Bestand sollte aus den gegebenen Flussgrößen errechnet werden. Von 12:00 bis 14:00 war die Regentonne leer, ab 14:00 stieg der Bestand konstant um 25/l pro Stunde. Um 18:00 war die Tonne schließlich mit 100 Litern gefüllt, und da dies genau dem Fassungsvermögen der Regentonne entspricht blieb auch der Bestand, trotz des anhaltenden Regens bei 100 Liter.

Von den 64 Testpersonen zeichneten immerhin 51 (=80%) den Graphen richtig ein. Nur vier Antworten waren komplett falsch. Die Textangabe wurde offensichtlich verstanden und die meisten zeichneten den Graphen richtig ein.

Aufgrund des hohen Anteils an korrekten Antworten und der recht kleinen Stichprobengröße verwundert es nicht, dass keinerlei signifikante Zusammenhänge zwischen dem Antwortverhalten auf Aufgabe F und den Variablen Geschlecht, Mathematiknote oder Selbsteinschätzung im Umgang mit Graphiken feststellen.

Man kann bei diesem Beispiel auch einen Vergleich mit der Aufgabe H ziehen. Grundsätzlich ging es bei beiden Fragen um das Einzeichnen eines Bestandes. Was sich aber schon unterschied, war die Angabe. Bei dem Beispiel F wurden die Flüsse verbal beschrieben, bei Aufgabe H jedoch waren aus grafisch gegebenen Flüssen der zugehörige Bestand zu ermitteln. Bei diesem Beispiel H konnte auch beim Posttest nur 50% aller Antworten korrekt, was darauf schließen lässt, dass die Teilnehmer trotz des Crashkurses noch immer einige Probleme hatten, grafisch gegebene Flussgrößen korrekt zu interpretieren.

5.2. Ergebnisse Aufgabe G

G) Die unterstehende Tabelle zeigt, wie viele Autos auf dem Uniparkplatz im Laufe eines Tages ankommen bzw. abfahren:

Stunde	An	Ab	Stunde	An	Ab
5-6	0	0	14-15	24	48
6-7	11	1	15-16	15	59
7-8	44	2	16-17	10	98
8-9	92	11	17-18	16	35
9-10	53	17	18-19	9	22
10-11	42	12	19-20	4	8
11-12	37	25	20-21	3	5
12-13	24	31	21-22	1	2
13-14	47	55	22-23	0	0

G1 Können Sie aus der Tabelle ermitteln, wann die meisten Autos auf dem Parkplatz stehen?
 Ja, um Uhr und zwar mit folgender Methode:

Nein, weil.....

G2 Wann sind die meisten Autos abgefahren?

Abbildung 5-2: Aufgabe G - Angabe

Die Story von Aufgabe G entspricht genau der von Aufgabe C beim Vortest. Der einzige Unterschied bestand darin, dass die Angabe bei Aufgabe G tabellarisch statt durch einen Graphen erfolgte. Außerdem sollte die Vorgangsweise, wie die Versuchspersonen auf den größten Bestand kommt bzw. warum dieser nicht zu ermitteln ist, gekommen sind, begründet werden.

Die meisten Fahrzeuge standen um 12:00 Uhr am Uni-Parkplatz (Frage G1). Die Argumentation läuft wie beim Vortest: bis 12:00 Uhr kommen jede Stunde mehr Autos an als abfahren und somit ist die Anzahl der Autos am Ende jeder Stunde größer als zu Beginn. Nach 12:00 ist es genau umgekehrt: in jeder Stunde fahren mehr Autos ab als ankommen. Wie beim Vortest wurde jedoch jede Zeitangabe zwischen 11 und 13 Uhr als richtig gewertet.

Insgesamt 51 Personen beantworteten diese Frage richtig. Weitere neun Personen sahen den höchsten Punkt der Ankünfte als den richtigen Zeitpunkt an, was wiederum darauf hinweist, dass diese Personen den größten Zufluss als maximalen Bestand annahmen.

Die richtige Begründung ist, dass bis 12:00 jede Stunde mehr Autos angekommen sind wie abgefahren und das es ab diesem Zeitpunkt umgekehrt ist. Dies war auch die meist genannte Begründung der Versuchspersonen.

Von den 51 Personen, welche den Zeitpunkt richtig erkannt hatten, gaben 27 diese Begründung an. Acht Personen gaben an, den Zeitpunkt durch Addieren und Subtrahieren ermittelt zu haben. Der Schnittpunkt der Flussgrößen wurde von drei Personen als Grund angegeben und die restlichen 13 Personen gaben leider keine Begründung an.

Sehr interessant sind die Probleme der drei Personen denen es unmöglich erschien diese Frage zu beantworten. Zwei davon gaben als Begründung den fehlenden Bestand an Autos vom Vortag an. Jedoch käme man auch zur gleichen Lösung, wenn man einen Bestand gegeben hätte. Man könnte jede Zahl als Bestand annehmen, das Ergebnis würde gleich bleiben, da es in diesem Fall der Nettozufluss wichtig ist. Eine Testperson war der Meinung, dass Flussgrößen gegeben sind und keine Diagramm. Diese Testperson hatte das graphische Beispiel vom Vortest jedoch richtig gelöst, woraus sich schließen lässt, dass sie einfach mit der Tabellenform überfordert war, und nicht erkannte, wie man in diesem Fall den "Schnittpunkt" bestimmt.

Bei der Aufgabe G2, die Frage nach der größten Abnahme der Autos, gaben sogar 59 Personen die richtige Antwort. Weitere 2 Personen irrten sich um plus/minus einen Tag. Die verbleibenden 3 Personen haben beim gesamten Beispiel überhaupt keine Antworten gegeben.

5.3. Ergebnisse Aufgabe H

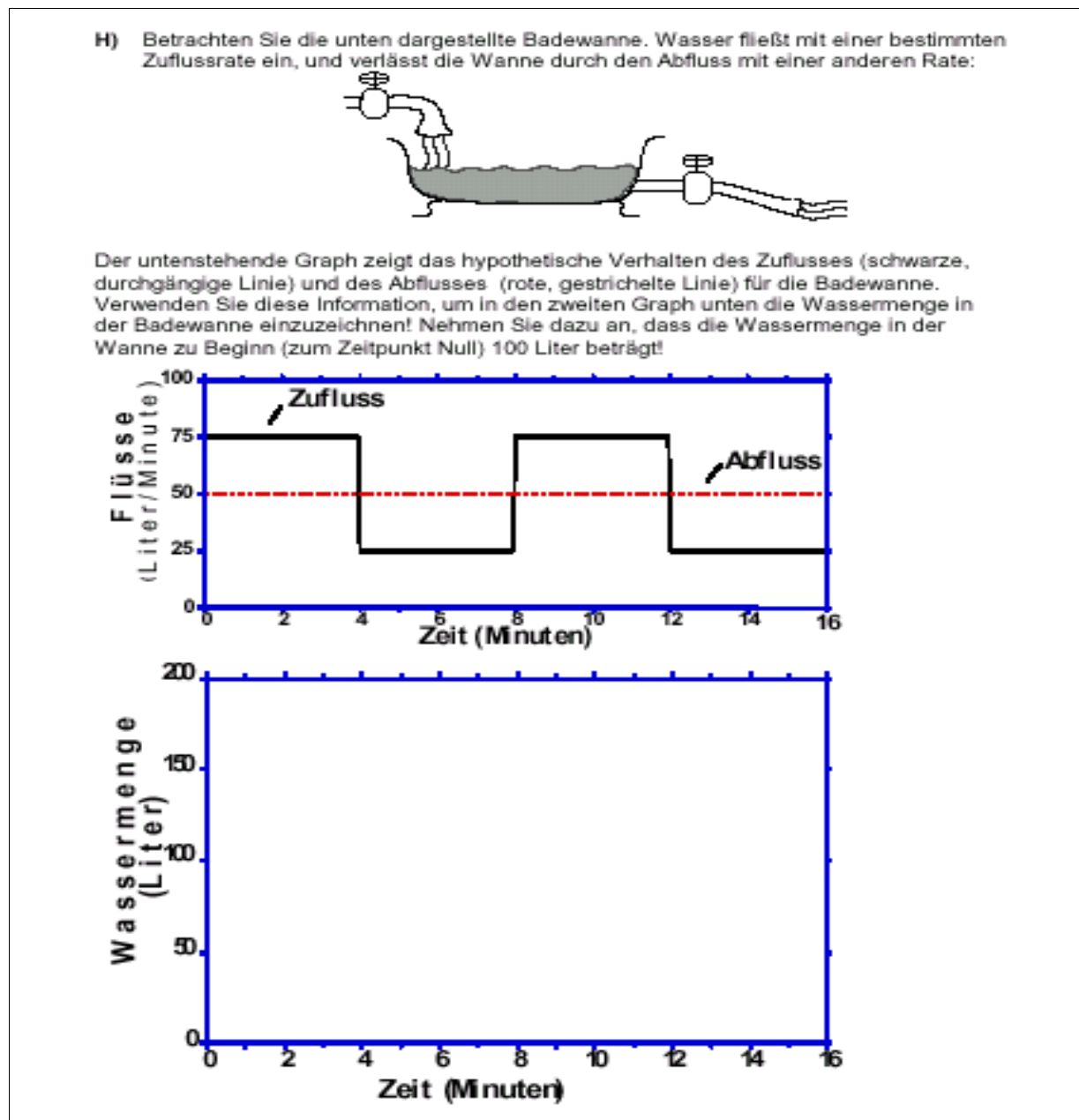


Abbildung 5-3: Aufgabe H - Angabe

Auch diese Aufgabe ist fast ident mit der Aufgabe D des Vortestes. Der einzige Unterschied ist, dass die Aufgabe H exakt in der Einkleidung, wie sie in der "Bathtub-

Aufgabe" von Sweeney/Sterman (2000) gestellt wurde. Die Aufgabe bestand darin, aus den gegebenen Zu- und Abflüssen die Wassermenge in der Badewanne zu ermitteln⁵.

Von jeder Lösung wurde wie beim Vortest beurteilt, inwieweit sie folgenden Kriterien genügt:

- Ist der Zufluss größer als der Abfluss, steigt die Wassermenge
- Ist der Abfluss größer als der Zufluss, sinkt die Wassermenge
- Der Graph soll keine diskontinuierlichen Sprünge enthalten (er ist stetig)
- Die Extremwerte des Graphen treten dann auf, wenn der Nettofluss gleich null ist
- In jedem Abschnitt ist der Nettofluss konstant, der Bestand muss linear steigen oder fallen
- Die Bestandsmaxima liegen bei 200 Litern nach 4 bzw. 12 Minuten, die Minima bei 100 Litern nach 8 bzw. 16 Minuten.

Vorweg kann man sagen, dass dieses Beispiel unter allen Aufgaben des Nachtests die meisten Probleme bereitete. Das Einzeichnen einer Zickzackkurve erschien vielen nicht plausibel. Wichtig war es, in diesem Beispiel zu erkennen, dass es sich hier um eine Bestandskurve handelte, die einzuzeichnen war.

Insgesamt 20 Personen zeichneten die korrekte Lösung ein, sie erfüllten somit alle oben genannten Kriterien für die richtige Einstufung der Graphik D.

Weitere 20 Personen erkannten immerhin den kontinuierlichen Zickzacklinienverlauf korrekt und zeichneten diesen fast richtig ein. Acht Teilnehmer erkannten, dass es sich zumindest um einen Anstieg und einen Abfall handelte, zeichneten dies aber sprunghaft und ähnlich wie in der Angabe ein. Sechs Teilnehmer hielten horizontale Geraden für die richtige Lösung. Die verbleibenden zehn Teilnehmer gaben keine Lösung an.

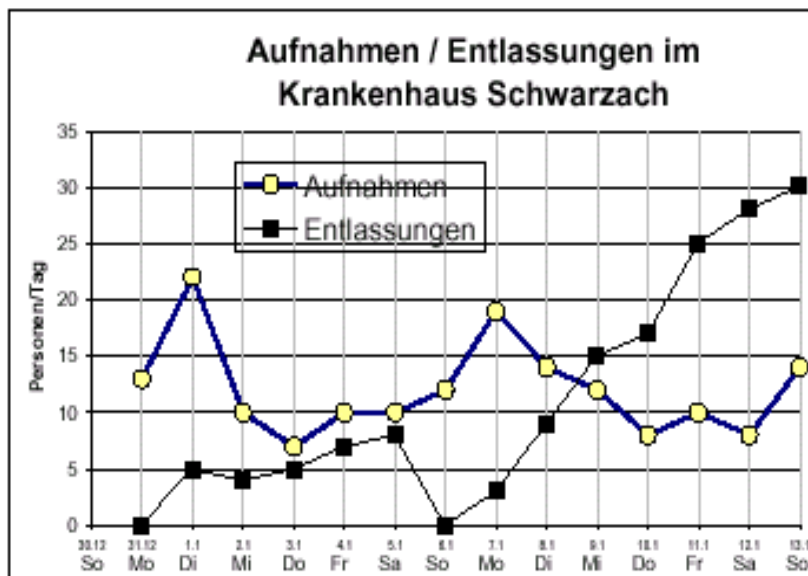
Auch hier konnte zwischen dem Antwortverhalten auf Frage H und den Variablen Geschlecht, Selbsteinschätzung im Lesen von Graphiken, Mathematiknote usw.

⁵ Sachlogisch erscheint die Einkleidung als Zu- und Abfluss einer Badewanne nicht so realistisch wie die Einkleidung mit dem Wasserausgleichsbehälter: Wer lässt schon in seiner Badewanne den Abfluss

keinerlei signifikanter Zusammenhang gefunden werden. Man kann lediglich einen leichten Trend feststellen, dass kann man sagen, dass die weiblichen Testpersonen etwas schlechter abgeschnitten haben als die männlichen VPn.

5.4. Ergebnisse Aufgabe K

- K) In der Ski-Saison werden viele Verletzte aus der Umgebung in das Krankenhaus Schwarzach-St.Veit eingeliefert. An den Hauptschitagen müssen die meisten neuen Patienten aufgenommen werden. Entlassungen finden hauptsächlich an Wochentagen statt. Durchschnittlich bleiben die Patienten 5-12 Tage in stationärer Behandlung. Die folgende Statistik zeigt für zwei Wochen für jeden Tag die Zahl der neu aufgenommenen und entlassenen Patienten:



Alle folgenden Fragen beziehen sich auf den in der Grafik erfassten Zeitraum!

- K1 An welchem Tag waren die meisten Patienten im Spital?
Tag: _____ bzw. lässt sich nicht beantworten, weil...
- K2 Wie sind Sie auf Ihre Antwort in Aufgabe K1 gekommen?
- K3 An welchem Tag haben die meisten Patienten das Spital verlassen?
Tag: _____ bzw. lässt sich nicht beantworten, weil...
- K4 An welchem Tag sind die meisten Betten frei geworden?
Tag: _____ bzw. lässt sich nicht beantworten, weil...

Abbildung 5-4: Aufgabe K - Angabe

permanent offen und lässt gleichzeitig ununterbrochen Wasser einlaufen, und das im Vierminutentakt abwechselnd stark und schwach?

In dieser Aufgabe ging es darum, wie schon auch der analogen Aufgabe B im Vortest, den maximalen Bestand an Patienten, die meisten Entlassungen sowie die größte Anzahl an frei werdenden Betten zu ermitteln. Es musste also ein Bestand, ein Abfluss und ein Nettoabfluss erkannt werden. Weiters sollte die Antwort auf die Frage nach den maximalen Bestand an Patienten begründet werden. Der Unterschied zum Vortest bestand darin, dass die Antworten in Beispiel H aus einer Tabelle herausgelesen werden mussten, wobei es sich hier um eine Kurvengraphik handelt.

Die Antwort auf die Frage nach den Tag mit den meisten Patienten im Spittal wäre korrekterweise die Nacht von 08.01. auf 09.01. gewesen. Dies wurde von 42 Personen richtig erkannt. Es gab nur eine Person, die um einen Tag mit ihrer Antwort daneben lag, jedoch insgesamt 16 Personen, welche das Datum mit dem größtem Patientenzuwachs (Saldo Aufnahmen-Entlassungen) als die richtige Lösung ansahen.

Die richtige Begründung wäre gewesen, dass es bis 08.01.2001 jeden Tag mehr Aufnahmen als Entlassungen gibt und dass ab 09.01.2001 jeden Tag mehr Entlassungen als Aufnahmen zu verzeichnen sind. Diese Begründung wurde von 15 der 42 Personen gegeben, welche die vorige Frage auch richtig beantwortet hatten. Weitere 23 Personen begründeten ihre Antwort auf die vorige Frage mit dem Schnittpunkt der Kurven. Als wahrhaft amüsant kann man die Antwort einer Versuchsperson bezeichnen. Nur eine Person gab an, durch die Addition der Veränderungen auf die richtige Lösung gekommen zu sein.

Eine weitere beliebte (allerdings falsche) Begründung - immerhin von 14 Personen - war, dass die höchste Differenz zwischen Aufnahmen und Entlassungen die richtige Lösung bringt. Diese Versuchspersonen haben, wie sich schon erahnen lässt, auch die vorige Frage nicht korrekt beantwortet.

In der dritten Frage war es die Aufgabe herauszufinden, wann die meisten Patienten das Spital verlassen haben. Die richtige Antwort war der 13.1., was auch von 61 Personen richtig erkannt wurde. Diese Frage fiel den Testpersonen anscheinend am leichtesten, da es die Frage mit den meisten richtigen Antworten ist. Eine Person gab das Datum mit der größten Patientenabnahme, also den Saldo zwischen Aufnahmen

und Entlassungen als richtige Lösung an und kam somit auf den 12.01.01. Zwei weitere Teilnehmer hatten Ableseungenauigkeiten bzw. Schreibfehler zu verzeichnen (31.01.01).

5.5. Ergebnisse Aufgabe L

- L Herr Maier lässt um Punkt 18:05 Uhr Wasser in seine Badewanne einlaufen. Die Zuflussmenge beträgt konstant 20 Liter/min. Um exakt 18:07 Uhr bemerkt Herr Maier, dass der Abfluss der Wanne offen war und schließt den Abfluss, lässt aber den Wasserhahn weiter offen. Aus dem Abfluss fließen pro Minute konstant 15 Liter Wasser. Um exakt 18:11 schließt Herr Maier den Wasserzufluss und genießt bis 18:20 die gefüllte Wanne. Um exakt 18:20 öffnet er den Abfluss und lässt die gesamten 90 Liter Wasser ausrinnen.
- L1 Skizzieren Sie in den beiden Grafiken in etwa, wie sich der Zufluss bzw. der Abfluss in der Badewanne im Laufe der Zeit entwickeln!

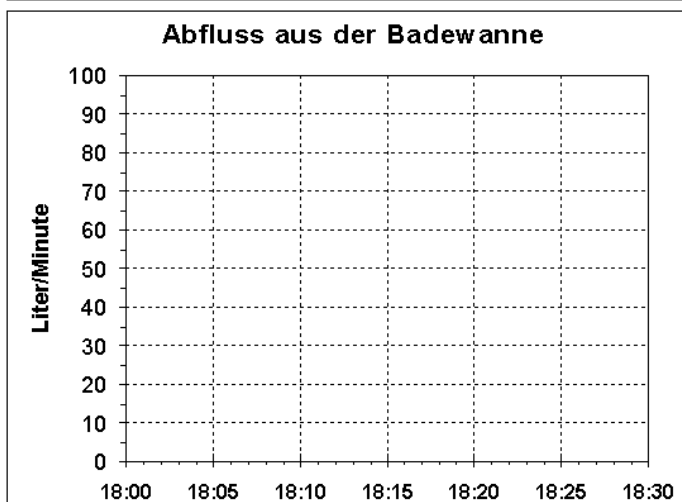
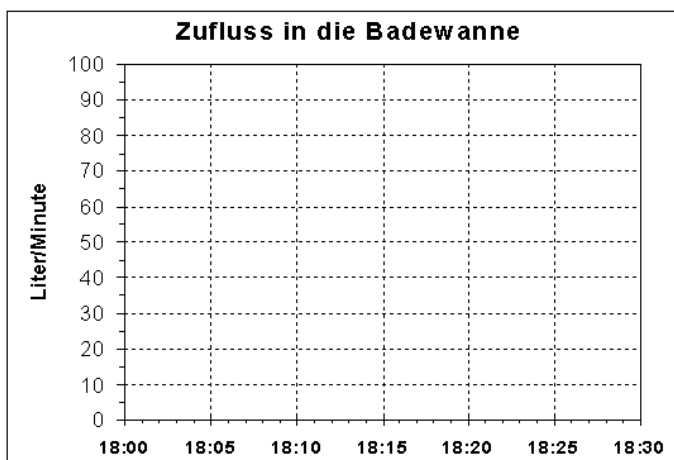


Abbildung 5-5: Aufgabe L - Angabe

Diese Aufgabe ähnelte dem Regentonnenbeispiel des Vortestes. Auch hier sollten die Teilnehmer einen konstanten Zufluss und einen Abfluss innerhalb der richtigen Zeitspannen erkennen.

Das Hauptproblem bei dieser Aufgabe hat sich im Vergleich zum Pretest kaum geändert. Weiterhin erkannten relativ wenige Versuchspersonen, dass hier Zu- und

Abflüsse gesucht waren und nicht der Bestand an Wasser in der Wanne, wie er von vielen VPn eingezeichnet wurde.

Der Zufluss wurde von 16 Teilnehmern korrekt eingezeichnet. Diese Personen erkannten einen Zufluss, der um 18:05 begann und um 18:11, als Herr Maier die Badewanne schloss, endete. Außerdem wurde von diesen VPn auch korrekt erkannt, dass es sich um einen maximalen Zufluss von konstant 20 Litern handelte. Während des Zeitraumes zwischen 18:00 und 18:05 sowie 18:11 und 18:30 betrug der Zufluss Null Liter.

Weitere vier Personen erkannten den Beginn und das Ende des Zuflusses, sowie auch die konstante Zuflussmenge von 20 Litern. Zwei Testpersonen erkannten den richtigen Beginn des Zuflusses und auch die konstante Zuflussmenge, der Zufluss endete jedoch nicht um 18:11, sieben Teilnehmer jedoch nur mehr die richtige Zeit von 18:05 bis 18:11. Unter diesen gezeichneten Graphen waren Zuflussmengen bis zu 100 Liter zu finden.

Die Mehrheit der Teilnehmer, nämlich 24 erkannten überhaupt nur mehr den Beginn des Zuflusses. Ein Ende des Zuflusses oder gar eine konstante 20 Liter Zuflussmenge waren nicht zu finden. Die restlichen 11 Personen zeichneten Graphen, die eigentlich nur mehr als „kurios“ einzustufen sind. Unter diesen Graphen kann man, angefangen mit Kurven und Treppenfunktionen, sehr viele „kreative“ Lösungen finden.

Der Abfluss der Badewanne wurde von 18 Teilnehmern korrekt eingezeichnet. Dieser beginnt um 18:05, da Herr Maier den Abfluss offen vergisst und endet um 18:07, wenn der Abfluss geschlossen wird. In der nächsten Zeit gibt es keinen Abfluss, bis schließlich Herr Maier um 18:20 den Abfluss wieder öffnet und die gesamten 90 Liter, mit einem konstanten Abfluss von 15 Liter pro Minute wieder ausrinnen lässt. Die Badewanne ist ab 18:26 wieder leer.

Die verschiedenen Varianten von Antworten der Versuchspersonen ähneln sehr stark den Problemen, welche die Teilnehmer auch mit dem Zufluss hatten. Dies lässt

wiederum darauf schließen, dass das Hauptproblem bei dieser Aufgabe, in der Unterscheidung zwischen Bestand und Flussgrößen liegt.

Zwischen dem Geschlecht der Teilnehmer und der Richtigkeit der Antworten besteht bei beiden Teilaufgaben kein Zusammenhang.

5.6. Zusammenfassung

Zum Schluss wurden auch noch für den Posttest die Gesamtzahl der richtigen Antworten, welche von jeder Person abgegeben wurden untersucht. Dazu teilten wir den Test in 11 Teilfragen, die es zu beantworten galt. Das unten angefügte Diagramm stellt das Ergebnis dar. Nur eine Person erreichte das Maximum an richtigen Antworten, jedoch gab auch niemand einen Test mit weniger als drei korrekten Lösungen ab, was im Gegensatz zum Pretest jedoch der Fall war. Der Median lag beim Posttest bei 7 richtigen Antworten (von 11), beim Pretest befindet sich dieser Wert bei 6 von 13, was man doch als eine spürbare Verbesserung in der Performance deuten kann. Wie schon im Pretest, kann man auch im Posttest keine Zusammenhänge zwischen Alter, Geschlecht, Maturanote und Selbsteinschätzung zum Graphikverständnis feststellen.

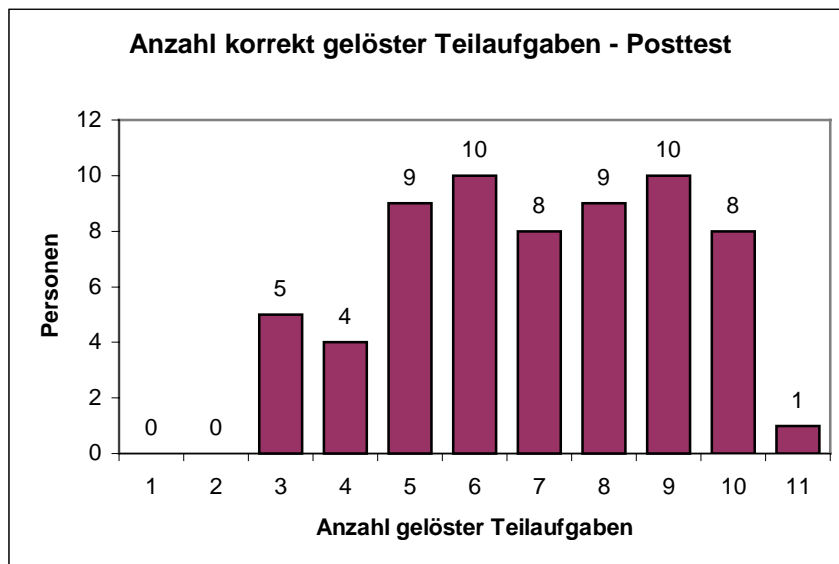


Abbildung 5-6 :Häufigkeitsverteilung der Antworten des Posttest

6. Vergleich Pretest – Posttest

(Melanie Zoltan und Martin Waiguny)

Nachdem wir nun die Ergebnisse des Pretests und des Posttests, erläutert haben, gehen wir nun auf die Zusammenhänge und Unterschiede in den Ergebnissen von Pre- und Posttest ein.

Dabei gehen wir beim Vortest von einer Grundbasis von 94 Personen und beim Nachtest von einer Basis von 64 Personen aus, d. h. Prozentangaben beziehen sich beim Pretest auf die Basis 100% = 94 Personen und Prozentangaben vom Posttest beziehen sich auf 100% = 64 Personen⁶. Man hätte als Alternative für den Vergleich auch alle Personen ausscheiden können, die nicht beide Tests gemacht haben.

Ergänzend konnten wir mit Hilfe der Codenamen die Leistungsentwicklung von einzelnen Teilnehmern direkt vergleichen. Hier zeigten sich interessante Verbesserungen; z. T. bei einzelnen VPn aber auch gewisse Verschlechterungen. Auf jeden Fall kann man bezüglich der Auswirkungen des Crashkurses einige vorsichtige Rückschlüsse ziehen.

Die beiden Tests waren vom Design her sehr ähnlich, die Aufgabengebiete waren gleich gewählt. Jedoch wurden die Beispiele in unterschiedlicher Verpackung (Stories) und leichter Variation dargestellt. Die folgende Tabelle gibt einen Überblick. (Den Vergleich der Beispiele der beiden Tests haben wir nach den zu erledigenden Haupt-Aufgaben gegliedert.)

⁶ Beim Vortest entspricht einer Person eine Veränderung von 1,064%, beim Nachtest schlägt sich eine Veränderung um eine Person prozentuell bereits mit über 1,5% nieder. Aus diesem Grund wurden die Prozentangaben auf ganze Prozent gerundet, weil hier Argumentationen mit Zehntelprozenten fehl am Platze wären.

Tabelle 6-1: Zusammenhang der Aufgaben von Vor- und Nachtest

Aufgabe	Vortest – Story	Nachtest – Story
Zu- und Abfluss zeichnen	Aufgabe A Regentonne	Aufgabe L Herr Maier badet
Bestände und Flüsse aus Tabelle lesen	Aufgabe B Krankenhaus	Aufgabe G Uniparkplatz
Bestände und Flüsse aus Graphik lesen	Aufgabe C Uniparkplatz	Aufgabe K Krankenhaus
Bestand von gegebener Graphik zeichnen	Aufgabe D Wasserausgleichsbehälter	Aufgabe H Badewanne*
Bestand von verbaler Angabe zeichnen	Aufgabe E Herr Maier badet	Aufgabe F Regentonne

* nach Sweeney/Sterman (2000)

6.1. Zu- und Abfluss zeichnen

Die zwei Aufgaben (A vom Vortest und L vom Nachtest) unterschieden sich kaum voneinander. Im Vortest handelt es sich um eine Regentonne mit einer beschränkten Füllmenge. Im Nachtest handelt es sich um eine Badewanne, mit Zu- und Abfluss an Wasser.

Bei beiden Beispielen waren zwei Aufgabe zu erfüllen. Einerseits musste der Zufluss eingezeichnet werden, andererseits auch der Abfluss in eine separate Graphik. Beide Graphiken waren mit der Zeit als Abszisse und der Flussmenge in Litern/Minute als Ordinate bereits vorskaliert.

Das Einzeichnen des Zuflusses bereitete beim Vortest den meisten Teilnehmern relativ große Probleme. Nur 5 von den 94 Testpersonen (=5%) konnten die Frage richtig beantworten. Beim Nachtest lässt sich eine leichte Verbesserung beobachten. Von den diesmal nur 64 Teilnehmer beantworteten 16 die Frage (=25%) richtig.

Erstaunlicherweise fiel es den Testpersonen leichter, den Abfluss korrekt einzuzeichnen. Hier gab es bereits beim Vortest 11 (=12%) richtige Antworten. Das Ergebnis beim Nachtest hat eine geringere Steigerung, als der Vergleich Zufluss beim Vor- und Nachtest, zu verzeichnen. Es gab 18 (=28%) richtige Antworten.

6.2. Werte aus Tabelle lesen – Werte aus Graphik ermitteln

Wir haben diese Aufgaben zusammengefaßt, da die Beispiele zwischen den beiden Tests nur in der äußeren Form der Aufgabenstellung variierten. Beim Pretest sollten die Teilnehmer die maximale Belegung eines Krankenhauses aus einer Tabelle ermitteln, beim Posttest gab es die gleiche Story mit den gleichen Zahlenwerten in Form einer graphischen Angabe. Genau umgekehrt verhielt es sich mit der Story vom Uniparkplatz. Zuerst war hier Graphikverständnis gefragt und dann gab es die für die meisten Teilnehmer wesentlich einfacher Tabellendarstellung.

6.2.1. Tabellendarstellung

Zuerst muss man ein Augenmerk auf die Ergebnisse der Tabellenbeispiele legen. Ein genauer Vergleich lässt sich hier nur zwischen den Fragen nach dem maximalen Bestand sowie nach dem größten Abfluss ziehen, da die Fragestellung leicht variierte.

Das Erkennen des maximalen Bestandes fiel sowohl beim Pretest als auch beim Posttest mehr als der Hälfte der Testpersonen leicht. Der Crashkurs bewirkte anscheinend dennoch eine Verbesserung der Fähigkeiten. Die relativen Anteile erhöhten sich von 60% beim Pretest auf 80% beim Posttest. Die Begründungen für die Antworten waren großteils sehr ähnlich obwohl beim Posttest nur mehr drei Personen angaben, dass diese Frage nicht beantwortbar wäre.

Leider lässt sich nicht ermitteln, ob die Teilnehmer das Maximum durch iteratives Saldieren aller Zu- und Abflüsse durch numerisches Rechnen herausgefunden haben, oder ob der Crashkurs dazu beigetragen hat, dass die Teilnehmer die Lösung anhand der Überlegung, dass in der ersten Phase an jedem Tag der Zufluss größer (und damit der Nettozufluss positiv) und in der zweiten Phase genau umgekehrt der Zufluss jeden Tag kleiner als der Abfluss (und damit der Nettofluss negativ) war, direkt aus der Tabelle ermitteln konnten.

Die Teilaufgabe den größten Abfluss betreffend wurde beim Posttest wie auch beim Pretest von den meisten Teilnehmern ohne größere Probleme gelöst. Auch die

relativen Anteile waren hier sehr ähnlich. Beim Pretest gaben bereits 90% Prozent den richtigen Wert an; beim Posttest gab es nur eine marginale Steigerung auf 92%.

6.2.2. Graphikdarstellung

Der Vergleich der Graphikbeispiele bringt wiederum erstaunliche Zahlen zum Vorschein. Hier zeigt sich eine eklatante Verbesserung. Hatten beim Vortest nur 17 Teilnehmer eine richtige Lösung abgegeben, so steigerte sich diese Zahl auf 42. Noch deutlicher fällt dies beim Vergleich der prozentuellen Zahlen auf. Wie haben hier eine gewaltige Steigerung von 18% auf 66%, also von gut einem Sechstel auf zwei Drittel aller VPn. Auch bei der Frage nach dem maximalen Zufluss gab es eine leichte Verbesserung von 88% auf 95% korrekter Antworten. Hier hatten die Teilnehmer bereits im Vortest relativ wenig Schwierigkeiten, so dass keine bedeutsamen Steigerungen mehr möglich waren. Etwas besser wird der Erfolg sichtbar, wenn man den Anteil der falschen Antworten betrachtet, der von 12% auf 5% zurückging.

Gerade bei diesem Beispiel sieht man, dass die entsprechenden Tipps des Crashkurses durchaus erfolgreich waren. Die häufigste Begründung, warum gerade diese an dieser Stelle das Maximum liegen müsse, war "Schnittpunkt der Kurven". Interessant wäre gewesen, ob die VPn auch im Falle mehrerer Schnittpunkte zwischen den Bestandskurven die Sache korrekt interpretiert hätten⁷.

6.2.3. Vergleich der Stories

Interessant ist allerdings der Vergleich zwischen den Stories. Die Ergebnisse, die hier vorliegen sind doch recht erstaunlich.

Beginnen wir mit der Krankenhausstory, die zuerst als Tabelle dargestellt wurde und für die wir im Posttest eine Graphikdarstellung wählten. Hier zeigt sich deutlich, dass

⁷ In diesem Falle repräsentiert jeder der Schnittpunkte der Flussgrößen ein (lokales) Extremum des resultierenden Bestandes. Wenn die Situation von "mehr Zuflüsse als Abflüsse" auf "mehr Abflüsse als Zuflüsse" wechselt, dann hat der Bestand ein Maximum erreicht, beim nächsten Wechsel von "mehr Abflüsse als Zuflüsse" wieder hin zu "mehr Zuflüsse als Abflüsse" hat der Bestand ein lokales Minimum usw.

der Crashkurs, vor allem beim Ablesen der Werte aus Graphiken, stark gefruchtet hat, wie folgende Tabelle, mit den ausschließlich als richtig gewerteten Ergebnissen, zeigt:

Tabelle 6-2 Korrekt gelöste Teilaufgaben der Krankenhausstory

Aufgabenstellung	Pretest (Tabellendarstellung) korrekt gelöst (n=94)	Posttest(Graphikdarstellung) korrekt gelöst (n=64)
maximaler Bestand	57 Teilnehmer = 60 %	42 Teilnehmer = 66 %
maximale Abfluss	84 Teilnehmer = 90 %	61 Teilnehmer = 95 %
größter Nettofluss	43 Teilnehmer = 46 %	28 Teilnehmer = 44 %

Insgesamt kann man nur leichte Verbesserungen der Versuchspersonen erkennen. Das Ablesen eines Nettoflusses aus einer Graphik fiel aber etwas schwerer als das direkte Ausrechnen desselben aus einer Tabelle. Dass in diesem Beispiel die Verbesserungen nicht stärker waren, liegt daran, dass die durch den Crashkurs erzielten Fortschritte durch die größere Schwierigkeit, einen graphisch gegebenen Fluss richtig zu interpretieren, größtenteils wieder kompensiert wurde.

Bei der Uniparkplatzstory lassen sich noch interessantere Resultate erkennen. Wie aus der folgenden Tabelle, in der die richtigen Ergebnisse aufgelistet sind, ersichtlich ist, gab es entscheidende Verbesserungen. Diese waren nicht nur auf Grund des Crashkurses zu erwarten, sondern auch deshalb, da die Tabellendarstellung bereits im Pretest für die meisten Teilnehmer einfacher zu interpretieren gewesen ist. Das ist auch der Grund, weshalb beim Nachtest bewusst, im Gegensatz zum Vortest, die Tabellendarstellung gewählt wurde.

Tabelle 6-3 korrekt gelöste Teilaufgaben Uni-Parkplatz-Story.

Aufgabenstellung	Pretest (Graphikdarstellung)	Posttest(Tabellendarstellung)
maximaler Bestand	17 Teilnehmer = 18%	51 Teilnehmer = 80%
maximaler Abfluss	83 Teilnehmer = 88%	59 Teilnehmer = 92%

Welche Schlüsse kann man nun aus diesen Daten ziehen? Zunächst einmal sieht man, dass beim Pretest der maximale Bestand bei einer tabellarisch gegebenen Aufgabe von wesentlich mehr Personen korrekt ermittelt wurde (60%) als bei grafisch gegebener Angabe, wo nur 18% der Antworten korrekt waren. Dies wird man wohl hauptsächlich damit erklären können, dass viele VPn bei tabellarisch gegebener Angabe in der Lage waren, sich die Entwicklung des Bestandes schlicht und einfach Tag für Tag durch Addition der Zuflüsse und entsprechender Subtraktion der Abflüsse auszurechnen. Im Zuge dieser Rechnung ergibt sich zwangsläufig das korrekte Maximum. Bei einer grafisch gegebenen Angabe scheint hingegen doch die "Verführungskraft der Kurve" bzw. das augenfällige Maximum der Zufluss-Kurve bei vielen VPn eine "magische Wirkung" auszuüben. Diese Wirkung bestätigt auch der Posttest: bei grafisch gegebener Angabe lösten zwei von drei VPn die Aufgabe korrekt, bei tabellarisch gegebener Angabe immerhin vier von fünf.

Die vorliegenden Ergebnisse erlauben noch eine weiter gehende Hypothese. Unabhängig davon, ob die Ausgangsdaten grafisch oder tabellarisch gegeben sind, immer führt die Überlegung: "Wenn der Zufluss in einem Zeitintervall größer ist als der Abfluss, dann ist der Bestand am Ende des Zeitintervalls höher als zu Beginn" recht unmittelbar zum Ergebnis, dass dort wo ein Überschuss an Zuflüssen in einen Überschuss an Abflüssen wechselt, der Bestand sein Maximum haben muss. Diese allgemeine Einsicht könnte man auch beim sequenziellen Saldieren der Zuflüsse und Abflüsse⁸ erhalten: immer wenn man einen Zufluss dazuaddiert und den entsprechenden Abfluss am selben Tag abzieht, dann ist der Bestand am Ende des Tages größer als zu Beginn des Tages. Dennoch scheint es so, als hätten die VPn, die dieses Rechenverfahren angewendet haben, diese Logik nicht erkannt – oder wenigstens waren sie nicht in der Lage, diese Logik auch im vergleichbaren Parallelspiel mit grafisch gegebener Angabe wiederzuerkennen.

Eine weitere Schlussfolgerung könnte sein, dass anscheinend das Prinzip der Akkumulation (oder mathematisch Integration, wenn man so will) zwar für numerisch gegebene Werte für die VPn intuitiv richtig umgesetzt wird, hingegen

⁸ Mit dieser Methode haben ja viele VPn bereits beim Vortest das Maximum des Bestandes korrekt ermittelt.

mit einer grafisch gegebenen Angabe keinerlei Integrations- bzw. Akkumulationsprozesse assoziiert werden. Sogar diejenigen Versuchspersonen, die korrekt erkannten, dass es auf die Abstände zwischen den beiden Kurven ankommt (also dem Saldo Zuflüsse – Abflüsse), blieben dann dennoch bei der Idee stecken, das Maximum des Bestandes wäre dort, wo die Differenz zwischen Zufluss und Abfluss am größten wäre. Tatsächlich ist dort nur die Nettoveränderung am größten.

Insgesamt kann man wohl feststellen, dass bei tabellarisch gegebenen Zu- und Abflüssen das fehlende Verständnis vom Unterschied zwischen Beständen und Flüssen sich weniger dramatisch auswirkt, weil die Vpn dies durch plausibles Rechnen kompensieren können. Bei grafisch gegebenen Flüssen schlägt das Problem, dass man Bestände und Flüsse nur schwer auseinander halten kann, voll durch.

Was hat nun der Crash-Kurs gebracht? Bei Ermittlung des Bestandes aus einer tabellarischen Angabe stieg der Anteil korrekter Antworten von 60% auf 80%; bei grafisch gegebenen Flüssen von 18% auf 66%. Vielleicht ist es hier angemessener, den Rückgang an falschen Antworten zu betrachten, weil für diejenigen Leute, die die Sache schon vor dem Crash-Kurs richtig gemacht haben, der Kurs keinen sichtbaren Effekt mehr erzielen kann. Bei den Aufgaben mit tabellarischer Angabe halbierte sich der Anteil falscher Antworten auf die Frage nach dem maximalen Bestand von 40% beim Pretest auf 20% beim Posttest. Bei den grafisch gegebenen Aufgaben verbesserte sich der Anteil der falschen Antworten sogar von 82% auf 34%! Insgesamt muss man jedoch vorsichtig sein, was die Interpretation der Pretest-Posttest-Vergleiche angeht. Immerhin haben ein Drittel der Personen, die den Vortest absolvierten, den Nachtest nicht mehr gemacht und überdies haben nicht alle Personen, die den Nachtest absolvierten, auch tatsächlich am Crash-Kurs teilgenommen. Zur Kontrolle wurden auch die Pretest-Leistungen derjenigen VPn untersucht, die beim Posttest nicht mehr dabei waren. Bemerkenswerterweise haben diese Personen beim Pre-Test überdurchschnittlich gut abgeschnitten, so dass wir hier die Wirksamkeit des Crash-Kurses eher konservativ geschätzt haben.

6.3. *Bestand von gegebener Graphik zeichnen*

Dieses Beispiel war im Nachtest (Aufgabe H) ident mit der Aufgabe D des Vortests. Es änderte sich nur die Story. Im Vortest handelte es sich um eher realistische Einkleidung eines Wasserausgleichsbehälters mit konstantem Abfluss und zwei Zufluss-Pumpen, während es sich im Nachtest um einen Zufluss in bzw. Abfluss aus einer Badewanne handelte. Beim Vortest gab es noch (ergänzende) Fragen zu den Pumpleistungen der einzelnen Pumpen, die hauptsächlich dazu dienten, zu klären, inwieweit die VPn die Angaben verstanden hatten.

Die Graphik im Pretest wurde von 15 der 94 Versuchspersonen (=16%) richtig eingezeichnet. Sehr viele Teilnehmer zeichneten annähernd den gleichen Graphen ein, wie den der Angabe. Bei diesem Beispiel lassen sich Auswirkungen des Crash-Kurses erkennen, denn beim Nachtest zeichneten immerhin 20 Personen, (31%) den richtigen Graphen ein. Damit wurde der Anteil korrekter Lösungen zwar verdoppelt, die z.T. massiven Probleme der Teilnehmer blieben aber dennoch ähnlich wie beim Vortest.

Die folgende Tabelle zeigt, wie sich die 64 Vpn, die sowohl den Vortest als auch den Nachtest absolviert haben, die beiden Aufgaben D (Vortest, VT) bzw. H (Nachtest, NT) bewältigten.

Tabelle 6-4: Vergleich Bathtub-Dynamics-Aufgabe VT und NT (Nur VP mit 2 Tests!)

	NT falsch	NT richtig	
VT falsch	38	15	53
VT richtig	6	5	11
	44	20	64

Die Zahl der richtig gelösten Aufgaben stieg vom Vortest zum Nachtest von 11 auf 20. Allerdings hat mehr als die Hälfte der 11 Personen, die beim Vortest die korrekte Grafik eingezeichnet haben, dies beim Nachtest NICHT mehr geschafft! Dieser Effekt ist ein Indiz dafür, dass die von Sweeney/Sterman (2000) stammende Einkleidung mit der Badewanne problematischer ist als die Einkleidung mit den Pumpen im Wasserausgleichsbehälter. Von den 53 Personen, die beim VT die Aufgabe nicht

korrekt gelöst haben, konnten immerhin 15 sie beim NT korrekt lösen. Hier könnte man einen Effekt des Crash-Kurses unterstellen.

6.4. Bestand von verbaler Angabe zeichnen

Die Teilnehmer mussten hier aus einer verbalen Angabe die richtige Bestandsgraphik ermitteln und einzeichnen. Beim Vortest war diese Aufgabe in der Einkleidung vom badenden Herrn Meier (Aufgabe E), beim Nachtest in der Einkleidung der sich füllenden Regentonne (Aufgabe F) gestellt. Wie man bereits aus den Einzelergebnissen der Tests erkennen kann, haben hier eher wenige Teilnehmer ein Problem mit der Interpretation der verbalen Angaben.

Die absolute Zahl an richtigen Lösungen waren beim Vortest wie auch beim Nachtest gleich: jeweils 51 Personen lösten die Aufgabe korrekt. Jedoch durch die geringere Teilnehmerzahl des Posttests war relative Anteil hier weitaus höher. Hier stehen 54% beim Vortest 80% beim Nachtest gegenüber, was einer deutlichen Verbesserung entspricht.

Die Vermutung, dass hier die selben 51 Personen beim zweiten Test teilgenommen haben liegt natürlich sehr nahe. Aber bei einem Vergleich der einzelnen VPn kommt man zum Ergebnis, dass 27 der 51 Vpn, die im zweiten Test eine korrekte Lösung von Aufgabe F lieferten, auch schon beim ersten Test die Aufgabe E korrekt lösten. Immerhin 23 der 51 Personen, die beim Vortest Beispiel E korrekt lösten, nahmen am Nachtest gar nicht mehr teil! Dies bedeutet, dass unter den 30 Abbrechern, die nur den ersten Test mitmachten und dann beim Posttest nicht mehr dabei waren, gleich 23 Personen waren, die Aufgabe E bereits korrekt lösten. Damit ist die These, dass vom ersten zum zweiten Test nur die „Versager“ aufgehört hätten, zumindest für das Beispielpaar E – F deutlichst widerlegt! Im ersten sowie im zweiten Test eine korrekte Lösung abgegeben haben. Umgekehrt haben natürlich auch 23 Personen, die beim Vortest Beispiel E falsch lösten, beim Nachtest Beispiel F richtig gelöst! Auch Verschlechterungen der Ergebnisse waren bei einzelnen Teilnehmer festzustellen.

Auch hier kann wiederum die Vermutung aufkommen, dass diese Teilnehmer nur Zufallstreffer tätigten. Dies mag auch an den fehlenden Sanktionierungs- und Honorierungsmöglichkeiten liegen. Auch zeigt sich, dass viele der Testpersonen anscheinend vor allem beim Pretest die Angabe nicht richtig durchdacht haben, da viele Graphen darauf hinweisen, dass nur ein schnelles Beantworten dieser Aufgabe wichtig war. Beim Posttest konnte anscheinend aufgrund des Crashkurses insgesamt doch ein etwas besseres Verständniss für diese Aufgabenstellung geschaffen werden.

6.5. Zusammenfassung

Aufgrund der in den vorherigen Kapitel Beispiele des Vor- und des Nachtests könnte man zu der Meinung kommen, dass sich keine sehr starken Verbesserungen bei den Teilnehmer bemerkbar machten. Durchschnittlich lag die Quote der Steigerung der Leistungen der Teilnehmer um die 20 Prozent.

Man darf jetzt allerdings nicht zum Schluss kommen, dass der Crashkurs eben nur diese vielleicht nicht überwältigend erscheinenden Verbesserungen gebracht hat. Wenn man vergleicht, welche Teilnehmer am Posttest teilgenommen haben, erkennt man, dass viele Versuchspersonen die beim Pretest bereits durchaus respektable Leistungen erbracht haben den Posttest, aus welchen Grund auch immer, ferngeblieben sind.

Ein sehr aussagekräftiger Beleg dafür wäre etwa das Regentonnenbeispiel des Pretests. Von den 5 Testpersonen, welche die korrekte Lösung ermittelten, nahm nur einer am Posttest teil. Ähnliches lässt sich auch bei anderen Beispielen beobachten. Insgesamt waren es also keineswegs nur die VPn mit den schwachen Leistungen beim Vortest, die dann nicht mehr zum Nachtest kamen, sondern die Leistungen der Ausscheider waren beim Pretest sogar z.T. ziemlich deutlich überdurchschnittlich.

Der Posttest hätte eventuell noch besser ausfallen können, falls diese Testpersonen teilgenommen hätten. Deshalb sind die durchaus guten Ergebnisse des Posttests sehr positiv zu bewerten, da hier die Testpersonen, welche größere Probleme mit Bestand- und Flussgrößen haben, vermehrt teilgenommen haben.

7. Ergänzende Interviews

(Eva Gratzner und Sandra Galli)

7.1. Ablauf

7.1.1. Auswahl der befragten Personen

Von den 94 beim Vortest befragten Personen wurden 20 für ergänzende Interviews ausgewählt. Dabei wurden Personen bevorzugt, die auf einzelne Fragen interessante Antworten gegeben haben. 13 Personen erklärten sich nach dem über die Code-Namen erfolgten Aufruf schließlich bereit, aus der Anonymität herauszukommen und sich für ein Interview zur Verfügung zu stellen. Von den 13 beim Vortest interviewten VPn konnten nur mehr 10 zum Nachtest interviewt werden. Nur die Hälfte dieser Personen war auch tatsächlich beim Crash-Kurs dabei. Die Ergebnisse der zum Nachtest interviewten Personen unterschieden sich deutlich, je nachdem, ob die jeweilige Person auch tatsächlich beim Crash-Kurs war oder nicht.

7.1.2. Durchführung der Interviews

Bei den Interviews wurden die VPn im wesentlichen danach gefragt, was sie sich beim Ausfüllen der Beispiele gedacht haben. Entscheidend war es, so weit wie möglich die Gedankengänge beim Ausfüllen der Fragebögen zu rekonstruieren. Es kam jedoch öfters vor, dass die VPn während des Interviews meinten, dass die beim Test gegebene Lösung nicht korrekt sei und eine andere Lösung angaben. Insofern war es schwierig, bei den Interviews das herauszuschälen, was die leitenden Gedanken während des Ausfüllens des Tests waren. Dazu kam noch z.T. ein beträchtlicher – natürlich vergeblicher! - Druck auf die Interviewer, zu erfahren, was nun die "richtige" Lösung sei.

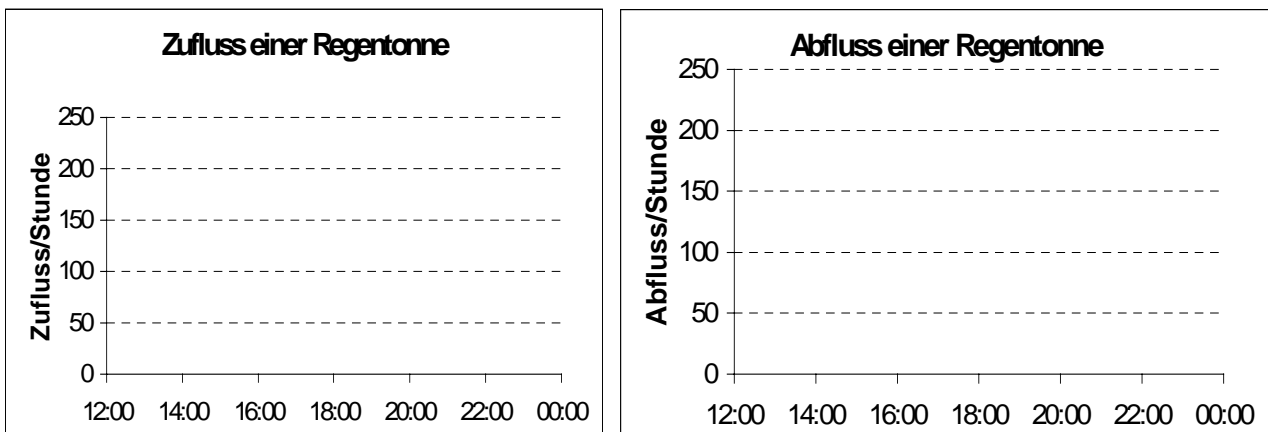
7.1.3. Dokumentation der Interviews

Die Gespräche wurden auf Tonband aufgezeichnet; besonders interessant erscheinende Passagen wurden später transkribiert. In diesem Bericht möchten wir zu jeder der fünf Aufgabenstellungen des Vortests eine knappe Gesamteinschätzung zur Aufgabe sowie einige besondere Interviewausschnitte präsentieren.

7.2. Interviews zum Vortest

7.2.1. Aufgabe A

- A) Eine Regentonne mit 100 l Fassungsvermögen ist um 12:00 Uhr leer. Um exakt 14:00 Uhr beginnt es leicht zu regnen und aus der Dachrinne füllt sich die Tonne mit 25l/Stunde Zufluss bis Mitternacht. Die Tonne besitzt keinen Abfluss, sondern wird so lange gefüllt, bis sie übergeht. Das übergehende Wasser wird als Abfluss gewertet. Zeichnen Sie den Zufluss und den Abfluss zwischen 12:00 Uhr und 00:00 Uhr in die Diagramme unten ein!



Im Kern sollte mit dieser Aufgabe überprüft werden, ob die VPn den Wasser**bestand** oder tatsächlich den Zu- und den Abfluss in die jeweiligen Diagramme eintragen. Eine Hilfestellung war dabei bereits die Tatsache, dass zwei Diagramme (und nicht nur ein Diagramm, in das Zu- und Abfluss gemeinsam einzutragen wären) angeboten wurden.

Zehn von den dreizehn interviewten Personen machten den Fehler, in das "Zufluss"-Diagramm nicht den Zufluss pro Stunde, sondern den kumulierten Zufluss, also den Bestand an Wasser in der Regentonne, einzutragen. Konkret wurde zwischen 14 und 18 Uhr ein linearer Anstieg von 0 auf 100 Litern eingetragen. Bei einigen VPn verläuft der Graph ab 18:00 horizontal, was andeutet, dass die Regentonne "voll" ist; bei anderen wuchs die Wassermenge linear weiter, bis um Mitternacht 250 Liter erreicht waren.

Auch der Abfluss wurde bei den meisten VPn als linear wachsender Graph eingetragen, wobei der Abfluss um 18:00 Uhr begann und um 00:00 Uhr 150 Liter erreichte. Insgesamt kann man sagen, dass die VPn die kumulierten Zuflüsse bzw.

Abflüsse graphisch dargestellt haben, also de facto jeweils den Bestand an in die Tonne geflossenem sowie aus der Tonne abgeflossenen Wasser. Diese Sichtweise bestätigte sich auch in den in den Interviews gegebenen Argumentationen.

Ein typisches Beispiel ist "Data 6", Jahrgang 1980, männlich. Seine Lösung sah folgend aus:

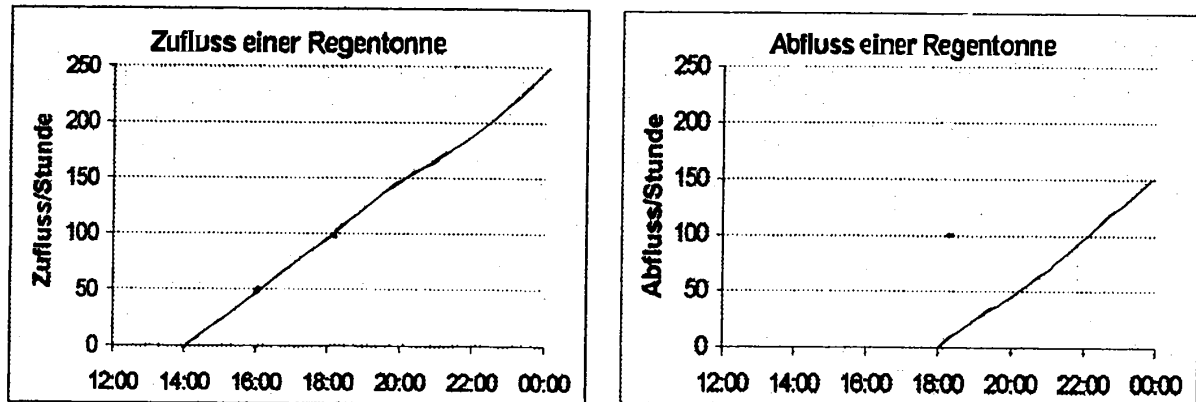


Abbildung 6-1: Aufgabe A - Lösung von "Data 6"

Ausschnitte aus dem Interview (D: Data 6; I: Interviewerin)

- 006 D OK, wir haben also zwei Graphiken: einmal den Zufluss in der Regentonne, einmal den Abfluss und zufließen tut von 14 Uhr bis null Uhr, deswegen habe ich da eingezeichnet, dass da von 14 Uhr bis null Uhr ein Zufluss stattfindet. Und es wird da immer mehr Wasser.
- 020 I Dass heisst, um Null Uhr hast Du wie viel drinnen?
- 021 D Um Null Uhr ist zugeflossen - es geht nicht darum, wie viel in der Tonne drinnen ist, sondern wie viel zugeflossen ist – um null Uhr ist insgesamt 250 Liter zugeflossen. Während die zweite Graphik nur den Abfluss darstellt. Um 18 Uhr ist die Tonne voll, und das heisst, ab 18 Uhr fließt Wasser ab und insgesamt fließen bis null Uhr 150 Liter Wasser ab.

Der Interviewausschnitt zeigt deutlich, dass Data 6 (korrekt) die kumulierten Zuflüsse bzw. Abflüsse graphisch dargestellt hat.

Ein zweiter Ausschnitt soll die eigenartige Lösung von Versuchsperson "Koweley" (Jahrgang 1978, m) erklären helfen. (K: Koweley; I: Interviewerin)

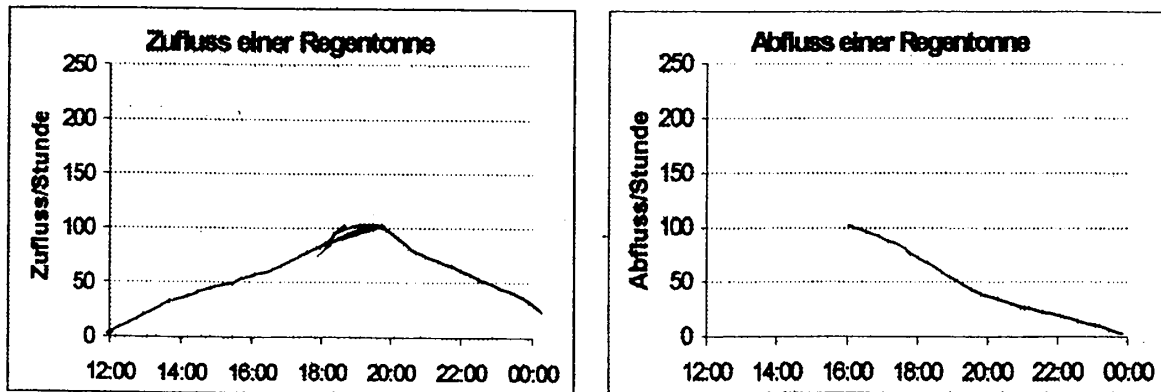


Abbildung 6-2: Aufgabe A - Lösung von "Koweley"

- 024 K** Also, ich habe mir das so gedacht, dass der Zufluss um 12 Uhr beginnt und der Höhepunkt ist dann ca. zwischen 18 und 20 Uhr. Es fällt dann stetig wieder bis dann um 0 Uhr das ganze wieder fertig ist.
- I** Und beim Abfluss?
- K** Beim Abfluss – aha, das übergehende Wasser wird als Abfluss gewertet – dh ab 16 Uhr geht die Tonne dann über und ab 16 Uhr kommt es dann stetig zu einem Abfluss, der dann auch um ca. 0 Uhr endet.
- 042 I** Mich würde jetzt interessieren warum der Abfluss bei 100 beginnt.
- K** Weil ich mir gedacht habe, dass die Regentonne bei 100 l Fassungsvermögen voll ist und ab diesem Zeitpunkt geht sie dann über (...)
- I** Wie würde man es jetzt interpretieren, dass der Abfluss um Mitternacht bei Null ist?
- K** Das ist jetzt insofern ein Gedankenfehler, weil ich einfach nur gedacht habe, dass ich den Abfluss fallend zeichnen sollte. (...)
- 064 I** Und wie bist Du hier auf den Zeitpunkt zwischen 18 und 20 Uhr gekommen, wo jetzt die steigende in die fallende Kurve übergeht?
- K** Aha, der Zufluss beginnt um 14 Uhr, dh die Tonne muss genau 4 Std. später voll sein – also um 18 Uhr. Ich habe es am Anfang auf der Zeichnung falsch eingezeichnet.

Bemerkenswert ist die Meinung von Koweley, dass ein Abfluss einen sinkenden Funktionsgraphen bedeute. Indirekt ist damit auch impliziert, dass Koweley hier von Beständen ausgeht, die durch einen Abfluss weniger werden. Was die Details des Zuflusses angeht (wann der Zufluss beginnt, wann er endet), so scheint es Koweley beim Zeichnen des Graphen eher locker genommen zu haben. Erst während des Interviews kam er drauf, dass man genau sagen kann, dass die Tonne um 18 Uhr voll ist.

7.2.2. Aufgabe B

- B)** In der Schi-Saison werden viele Verletzte aus der Umgebung in das Krankenhaus Schwarzach-St.Veit eingeliefert. An den Hauptschitagen müssen die meisten neuen Patienten aufgenommen werden. Entlassungen finden hauptsächlich an Wochentagen statt. Durchschnittlich bleiben die Patienten 5-12 Tage in stationärer Behandlung. Die folgende Statistik zeigt für zwei Wochen für jeden Tag die Zahl der neu aufgenommenen

und entlassenen Patienten:

	Datum	Aufnahmen	Entlassungen	Datum	Aufnahmen	Entlassungen
So	31.12.2000	13	0	07.01.2001	19	3
Mo	01.01.2001	18	0	08.01.2001	11	9
Di	02.01.2001	10	4	09.01.2001	12	15
Mi	03.01.2001	7	5	10.01.2001	8	17
Do	04.01.2001	10	7	11.01.2001	10	25
Fr	05.01.2001	10	8	12.01.2001	8	28
Sa	06.01.2001	12	0	13.01.2001	14	30

Alle folgenden Fragen beziehen sich auf den in der Tabelle erfassten Zeitraum!

- B1 An welchem Tag waren die meisten Patienten im Spital?
Tag: _____ bzw. O lässt sich nicht beantworten, weil....
- B2 Wie sind Sie auf Ihre Antwort in Aufgabe B1 gekommen?
- B3 An welchem Tag haben die meisten Patienten das Spital verlassen?
Tag: _____ bzw. O lässt sich nicht beantworten, weil....
- B4 An welchem Tag sind die meisten Betten frei geworden?
Tag: _____ bzw. O lässt sich nicht beantworten, weil....

Die meisten VPn lösten die Frage B1 durch direktes Ausrechnen, indem die Zahl der Aufnahmen und Entlassungen fortlaufend saldiert wurden und kamen damit zur korrekten Antwort 8. / 9. Jänner.

VP "14-10" (1981, m) nannte den 1.1. als den Tag mit den meisten Patienten im Spital und begründete dies folgend:

- 129 14** Grundsätzlich ist es einmal so, da ich nicht alle Daten vorliegen habe, wie viele Leute in diesem Krankenhaus da sind – oder nur die Aufnahmen und Entlassungen gegeben habe – kann ich das eigentlich nicht feststellen, jedoch habe ich mir gedacht, dass sich dies auf die Aufnahmen und Entlassungen bezieht, und eben am Montag, den 1.1. 18 Aufnahmen waren und keine Entlassungen.
- 139 I** Also hast Du dir da praktisch die Differenz ausgerechnet...
- 14** richtig, ja; also da wäre noch in Frage gekommen der 7.1., da waren 19 Aufnahmen, aber 3 Entlassungen, also waren meines Erachtens am 1.1. waren die meisten Patienten im Spital – also, was die Aufnahmen und Entlassungen angeht.

Der Ausschnitt zeigt deutlich, dass "14-10" die Differenz zwischen Aufnahmen und Entlassungen eines Tages heranzieht, um das Maximum der Patienten zu ermitteln. Damit wird de facto ein Nettofluss mit einem Bestand verwechselt.

Die Fragen B3 und B4 wurden von den meisten VPn als gleichartig angesehen und der 13.1. als (für B3 korrekte) Antwort gegeben. Ob mit "frei gewordenen" Betten der Saldo zwischen Entlassungen und Neuaufnahmen oder ebenfalls nur die Entlassungen gemeint ist, ist mehr ein semantisches Problem. Alle interviewten VPn konnten jedenfalls ihre Antwort plausibel begründen.

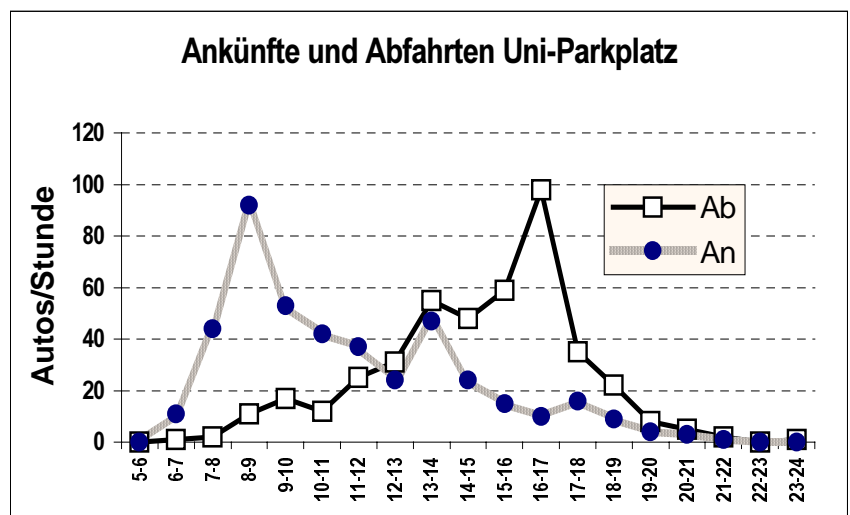
In manchen Fällen argumentierten die VPn damit, dass die Frage B1 nicht beantwortbar wäre und begründeten dies damit, dass die tatsächliche Zahl der Patienten im Krankenhaus nicht bekannt sei, weil nur Aufnahmen und Entlassungen wiedergegeben werden.

7.2.3. Aufgabe C

C) Die nebenstehende Graphik zeigt, wie viele Autos auf dem Uniparkplatz im Laufe eines Tages ankommen bzw. abfahren (vor 5 Uhr früh ist der Parkplatz leer):

C1 Können Sie aus der Graphik sicher ermitteln, wann die meisten Autos auf dem Parkplatz stehen?

Ja, um Uhr und zwar mit folgender Methode:



Nein, weil.....

C3 Wann sind die meisten Autos abgefahren?

Auch bei diesem Beispiel wurde die Zahl der innerhalb einer Stunde ankommenden mit der Zahl der am Parkplatz stehenden PKW verwechselt und angegeben, dass in der Zeit von 8-9 Uhr (oder um 9 Uhr) sich die meisten Autos am Parkplatz befinden.

Eine Variante dieser Argumentation bietet VP "FM3" (1981, w):

063 I Du hast gesagt: "Die meisten Autos sind am Parkplatz zwischen 8 und 9 Uhr". Wie bist du auf das gekommen?

065 F Ich habe mir angeschaut, wann am meisten Autos angekommen sind und dann im Vergleich dazu, wo relativ wenige abgefahren sind und dadurch, dass da eine Spitze ist bei den Ankünften und aber doch die Anzahl der Abfahrenden relativ gering ist – habe ich dann angenommen, dass die Differenz da doch relativ am größten ist.

VP FM3 argumentiert also mit der Differenz zwischen ankommenden und abfahrenden Autos, also mit einem Nettofluss (der natürlich ebenfalls etwas anderes ist als der Bestand an Autos.) Dass man bei dieser Aufgabe auch noch auf eine andere Art Bestand und Flüsse mit einander verwechseln kann, belegt das Beispiel von VP Mercedes (1975, m):

144 I Wann stehen die meisten Autos auf dem Parkplatz?

M Zwischen 16 und 17 Uhr weil da am meisten Autos abfahren. Also müssen auch am meisten da sein wenn am meisten wegfahren. Da fahren am meisten Leute weg; also müssen auch vorher am meisten Leute dort gestanden sein.

Auch das Problem, dass keine Ausgangszahl an Autos angegeben ist, machte einigen VPn Kopfzerbrechen. Bei VP "Gartner" (1976, m) zeigt sich wiederum deutlich, dass die VP bei ihrer Argumentation eigentlich eine Bestandskurve im Kopf hatte:

219 I Und zur Frage C1 – wann sind jetzt die meisten Autos am Parkplatz gestanden?

G Hier ist das gleiche Problem wie bei der Frage C2 – es müssten vorher null Autos am Parkplatz gestanden sein, dass ich es überhaupt ausrechnen kann.

234 I Was macht es für einen Unterschied ob am Anfang null Autos oder 50 gestanden sind?

G Ja ich habe einfach einen Sockelbetrag den ich hier miterfasse. Das würde das Ganze linear nach oben verschieben.

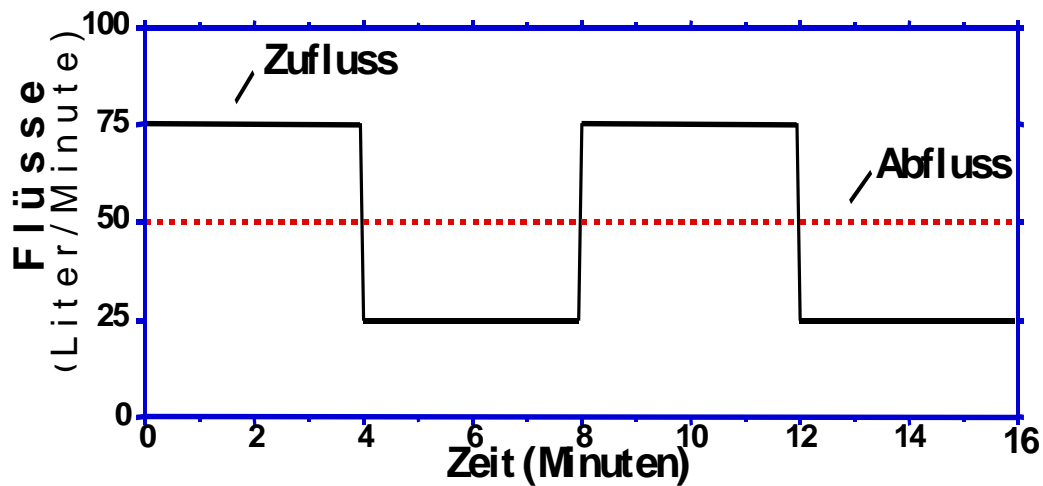
260 I Und wenn du die Graphik nur schnell anschaust, wann ungefähr würden die meisten Autos am Parkplatz stehen?

G Wenn ganz viele zufahren und ganz wenig wegfahren. Also das müsste ca. um neun Uhr sein.

Gartner's Argumentation "Das würde das Ganze linear nach oben verschieben" wäre dann korrekt, wenn es sich um eine Kurve der Bestände handeln würde. Die Zahl der Ankünfte bzw. Abfahrten ändert sich durch einen anderen Ausgangsbestand natürlich nicht. In 260 argumentiert Gartner wieder genauso wie FM3, dass die Differenz zwischen Ankünften und Abfahrten das relevante Kriterium für die maximale Zahl an Autos am Parkplatz wäre.

7.2.4. Aufgabe D

D) Ein Wasserausgleichsbehälter soll gleichmäßig Wasser 50 Liter Wasser pro Minute abgeben. Der Behälter wird von zwei Zufluspumpen gespeist. Die Dauer-Zufluspumpe liefert gleichmäßig eine gewisse Wassermenge pro Minute, Die zweite Zufluspumpe ist eine Kurzzeit-Hochleistungspumpe, die nur vier Minuten lang gleichmäßig Wasser zupumpt, dann aber für die nächsten vier Minuten abgeschaltet werden muss, damit der Brunnen, aus dem sie fördert, sich wieder auffüllen kann. Der untenstehende Graph zeigt den Abfluss (gestrichelt) und den Gesamt-Zufluss (durchgängige Linie) des Ausgleichsbehälters.



- D3 Tragen Sie in das untenstehende Diagramm ein, wie sich die Wassermenge im Ausgleichsbehälter in den betrachteten 16 Minuten entwickelt, wenn die im obigen Graphen gegebenen Zu- und Abflussverhältnisse herrschen! Nehmen Sie an, dass der Behälter zu Beginn 100 Liter enthält!

Einer der häufigsten Fehler bei diesem Beispiel war, dass im wesentlichen die Struktur der Eingabegraphik wiederholt wurde. Ein Beispiel dafür ist VP "14-10" (1981, m)

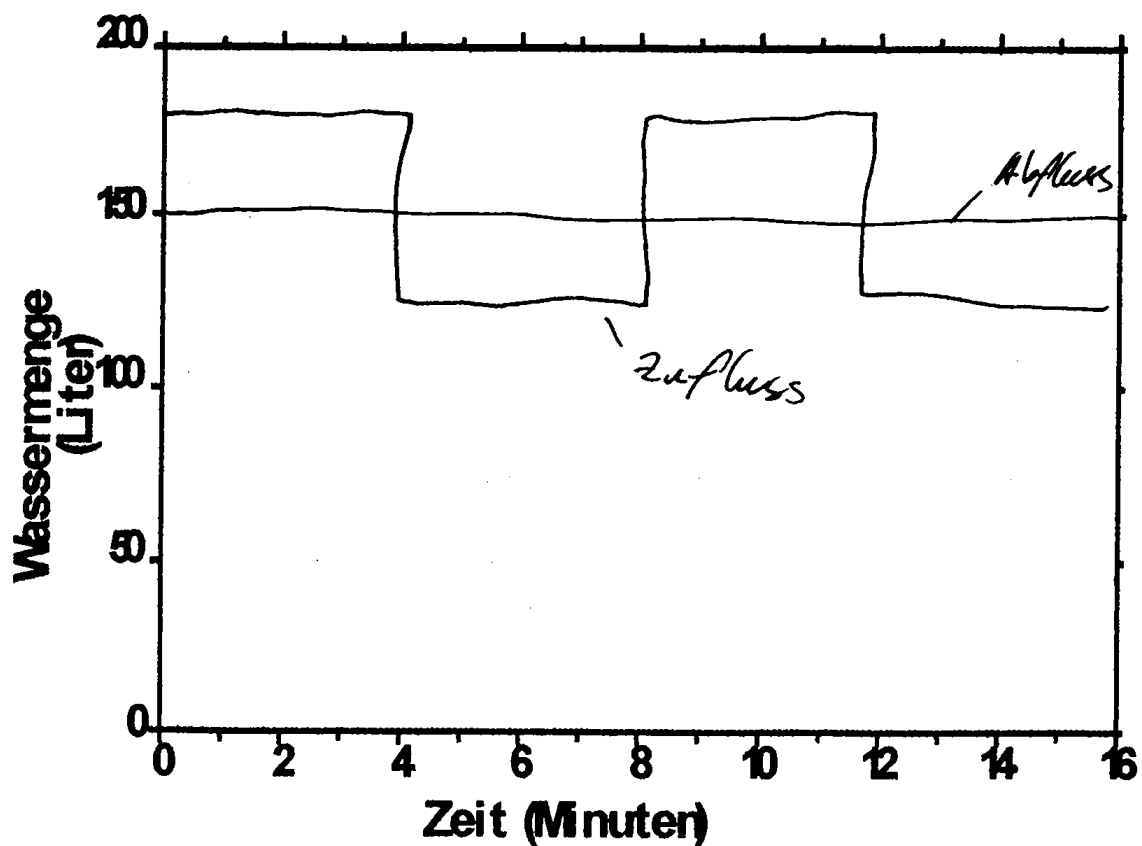


Abbildung 6-3: Aufgabe D - Lösung von "14-10"

284 14 Hier habe ich wiederum die Wassermenge gegeben, hab jetzt aber schon 100 Liter in diesem Behälter drinnen. Deshalb addiere ich – mehr oder weniger – diese 50 Liter von der oberen Graphik dazu und dadurch habe ich dann bei 150 Liter eine Gerade, die mir hier den Abfluss anzeigt

Im Hintergrund steht wiederum die Unfähigkeit, zwischen einer Graphik, die Bestandsgrößen und einer Graphik, die Flussgrößen wiedergibt, zu unterscheiden. Hinter dem von "14-10" gewählten Konzept, die zu Beginn vorhandenen 100 Liter zu addieren, steckt wiederum die Annahme, es handle sich beim Zufluss und Abfluss in der Angabe (wie dann auch in der zweiten Zeichnung) jeweils um Bestandsdaten.

7.2.5. Aufgabe E

E) Herr Maier lässt um Punkt 19:00 Uhr Wasser in seine Badewanne einlaufen. Die Zuflussmenge beträgt konstant 14 Liter/min. Um exakt 19:04 Uhr bemerkt Herr Maier, dass der Abfluss der Wanne offen war und schließt den Abfluss. Aus dem Abfluss fließen pro Minute konstant 9 Liter Wasser. Um exakt 19:09 schließt Herr Maier den Wasserzufluss und genießt bis 19:15 die gefüllte Wanne. Um exakt 19:15 öffnet er den Abfluss und lässt das gesamte Wasser ausrinnen.

Die meisten VPn zeichneten in etwa eine korrekte Graphik. Die untenstehende - fast richtige - Lösung stammt von VP "Dodo" (1981, w).

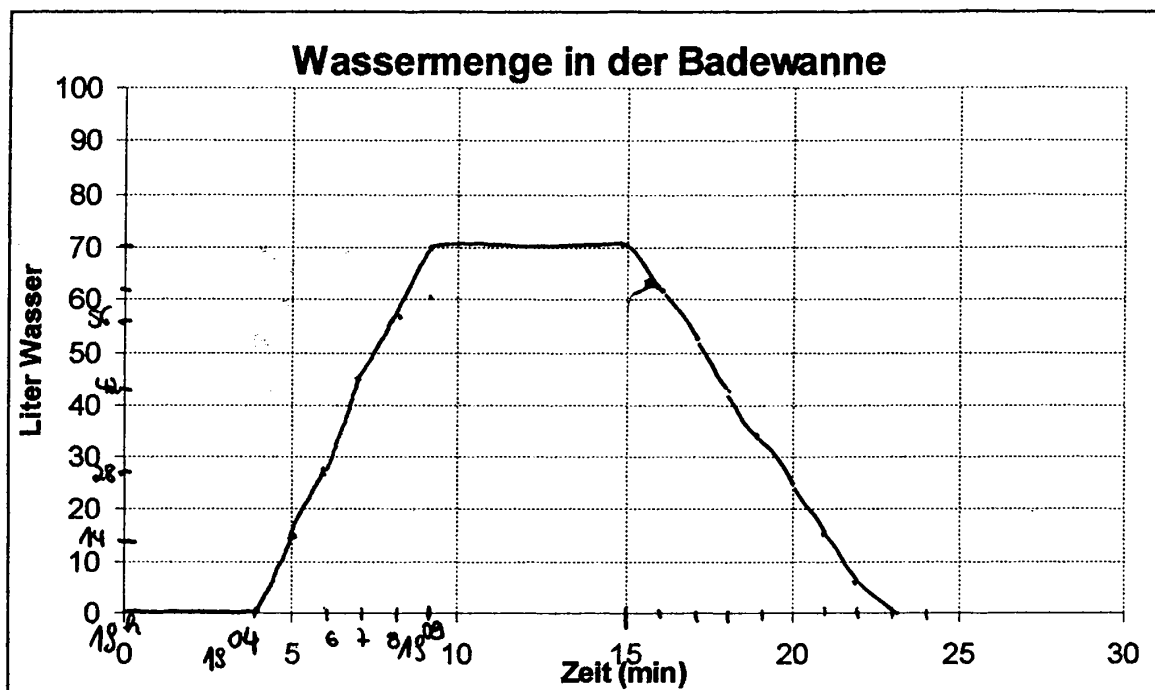


Abbildung 6-4: Aufgabe E - Lösung von "Dodo"

170 I Kannst du mir nun das Badewannen Beispiel erklären?

D Also der Hr. Maier lässt um 19 Uhr Wasser einlaufen. Ich habe die ersten 4 Minuten deshalb keinen Abfluss eingezeichnet, weil der Abfluss offen war. Erst um 19:04 habe ich einen Zufluss von 14 l/min eingezeichnet. Um 19:09 schließt er dann den Wasserzufluss, dh von

19:09 bis 19:15 bleibt der Wasserstand konstant bei 70 l. Um 19:15 ist dann ein Abfluss von 9 l/min. Um 19:23 ist die Badewanne dann wieder leer.

VP Simba (1981, w) erläutert ihren linear wachsenden Graphen folgend:

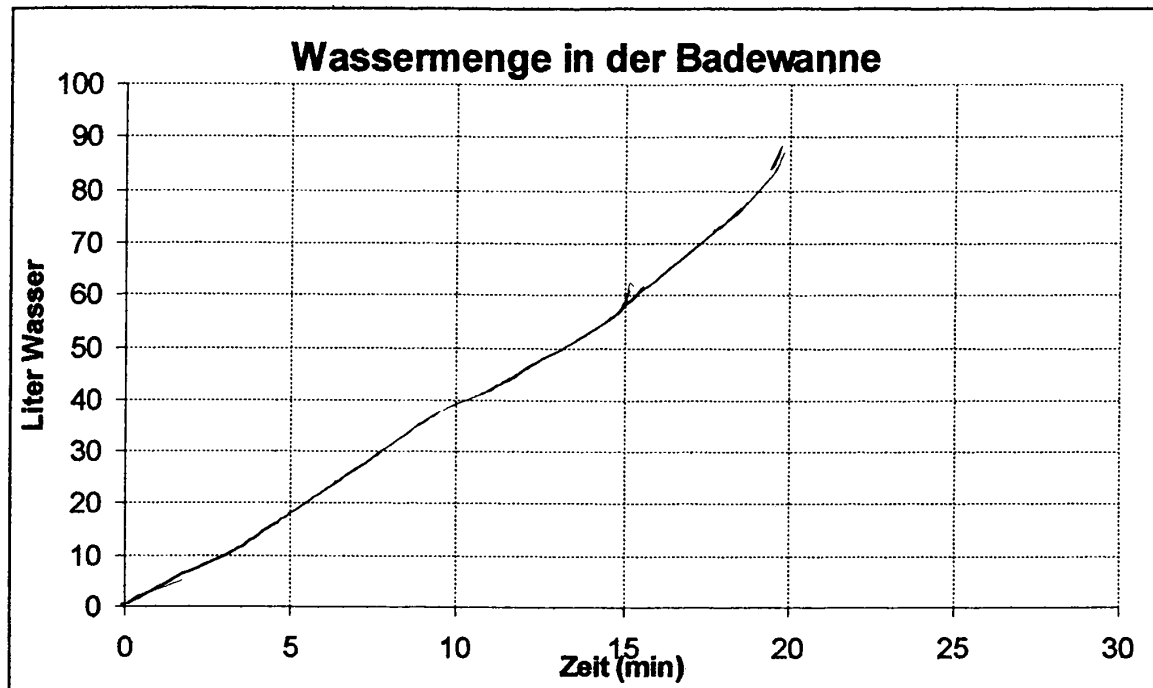


Abbildung 6-5: Aufgabe E - Lösung von "Simba"

- 183 I** Kannst du dich noch an dieses Beispiel erinnern und warum hast du hier eine kontinuierlich ansteigende Linie gezeichnet?
- S** Im Prinzip ist es ja so, dass wenn man Wasser in eine Wanne einlässt, steigt das Wasser kontinuierlich pro Minute, weil es wird nicht einmal schneller und einmal langsamer sinken. Und meiner Meinung nach ist es auch dem entsprechend zu zeichnen.

Das Interview macht deutlich, dass die VP mit dieser Aufgabe anscheinend überfordert war: Trotz nochmaliger Aufforderung, die Angabe zu überprüfen, machte Simba keinerlei Anstalten, die Informationen, wann was beim Füllen der Badewanne geschieht, auch irgendwie in den Graphen umzusetzen. Obendrein werden von ihr anscheinend die Wörter "steigen" und "sinken" leicht durcheinander gebracht, wie auch obiger Ausschnitt belegt. Insgesamt stiftete die Interviewpassage bei uns Interviewerinnen mehr Konfusion als sie Klärung der Denkvorgänge von Simba brachte.

7.2.6. Zusammenfassung der Interviews zum Vortest

In einigen Fällen konnten durch die Interviews auch falsche Ansätze der VPn plausibel erklärt werden. In mehreren Fällen jedoch hatten wir den Eindruck, dass sich einzelne VPn die Aufgaben ziemlich leicht gemacht hatten – und beim Vortest einfach rasch etwas hingezeichnet haben, ohne die Angabe überhaupt genau zu lesen. Dies äußerte sich darin, dass diese VPn auf die Frage, warum sie die betreffende Aufgabe gerade so und nicht anders gelöst hätten, zunächst meinten "Da kann ich mich nicht so genau erinnern" und dann erst langwierig die Angabe zu studieren begannen. Bisweilen verstiegen sich die VPn auch in krause Nebengedanken oder andere Strategien zur Verschleierung, dass Sie entweder mit der Aufgabenstellung gar nicht richtig vertraut oder damit auch überfordert waren. Wir haben in solchen Situationen allerdings keinen besonderen Druck auf die VPn ausgeübt, sondern versucht, das Gespräch in einem freundlichen Rahmen zu halten; auch um den Preis, dass gewisse Klärungen damit dann nicht mehr möglich waren.

7.3. Interviews zu den Nachttests

Beim Nachttest wurden im wesentlichen dieselben Aufgaben wie beim Vortest gegeben. Allerdings wurden alle Aufgaben leicht variiert. Die folgende Tabelle gibt einen Überblick

Story	Vortest	Nachttest
1 Regentonne	A: Zu- und Abfluss zeichnen	F: Bestand zu zeichnen
2 Krankenhaus	B: Tabelle gegeben	K: Graphik gegeben
3 Uni-Parkplatz	C: Graphik gegeben	G: Tabelle gegeben
4 Bathtub-Dynamics	D: Wasserausgleichsbehälter	H: Badewannenfüllung*
5 Herr Maier badet	E: Bestand zu zeichnen	L: Zu- und Abfluss zeichnen

* nach Sweeney/Sterman 2000

Bei der Regentonnen-Story 1 war beim Vortest Zu- und Abflüsse zu zeichnen; beim Nachttest hingegen war (mit der selben Angabe wie beim Vortest) der Bestand zu zeichnen. Umgekehrt war bei der Story 5 "Herr Maier badet" beim Vortest der Bestand an Wasser in der Badewanne, beim Nachttest hingegen der Zu- bzw. Abfluss zu zeichnen. Also entsprechen die Aufgaben A und L bzw. E und F von der

inhaltlichen Fragestellung einander, wobei die Stories 1 und 5 jeweils vertauscht wurden.

Analog gehören die Stories 2 (Krankenhaus) und 3 (Uni-Parkplatz) zusammen. In Aufgabe B beim Vortest wurden die Aufnahmen und Entlassungen als Tabelle gegeben. Analog dazu war beim Nachtest die Uni-Parkplatz-Story in Aufgabe G als Tabelle gegeben. Als Graphik gegeben war die Uni-Parkplatz-Story bei Aufgabe C beim Vortest und die Krankenhaus-Story in Aufgabe K beim Nachtest.

Die Bathtub-Dynamics-Story stammt ursprünglich von Sweeney und Sterman (2000). Die Aufgabe H beim Nachtest ist exakt die Übersetzung der Sweeney/Sterman-Aufgabe. Beim Vortest wurde in Aufgabe D eine andere Einkleidung (als Wasserausgleichsbehälter, der von zwei Pumpen gespeist wird) verwendet, die die abrupten Wechsel im Zufluss plausibler erklärt als in der Badewannen-Einkleidung.

Insgesamt werden also Aufgabe A mit Aufgabe L, Aufgabe B mit Aufgabe G, Aufgabe C mit Aufgabe K, Aufgabe D mit Aufgabe H und Aufgabe E mit Aufgabe F verglichen, wobei in keinem Fall die Stories exakt gleich geblieben sind.

7.3.1. Aufgabe K

VPn Gartner hat beim Vortest bei Beispiel C so argumentiert, dass man darauf schließen kann, dass er die Graphik Veränderungen mit einer Graphik der Bestände verwechselt. Beim Nachtest hat er bei Aufgabe K korrekt erkannt, dass die Zahl der Patienten dann am höchsten ist, wenn sich die beiden Kurven schneiden.



Alle folgenden Fragen beziehen sich auf den in der Grafik erfassten Zeitraum!

K1 An welchem Tag waren die meisten Patienten im Spital?
 Tag: 8.1 bzw. lässt sich nicht beantworten, weil....

K2 Wie sind Sie auf Ihre Antwort in Aufgabe K1 gekommen?
Da am 8.1 ist die Anzahl der Aufnahmen geringer als die Anzahl der Auskünfte \Rightarrow 5 Patienten im Spital sinkt.

K3 An welchem Tag haben die meisten Patienten das Spital verlassen?
 Tag: 1.1 bzw. lässt sich nicht beantworten, weil....

K4 An welchem Tag sind die meisten Betten frei geworden?
 Tag: 1.1 bzw. lässt sich nicht beantworten, weil....

Abbildung 6-6: Aufgabe K - Lösung von "Gartner"

7.3.2. Aufgabe L

Versuchsperson "Data 6" hat beim Vortest bei Aufgabe A statt der Flüsse jeweils den akkumulierten Zufluss (also den Bestand) an Wasser in der Regentonne bzw. den Bestand an ausfließendem Wasser eingezeichnet. Beim Nachtest hingegen gab er bei Aufgabe L die gefragten Flüsse (fast) korrekt wieder:

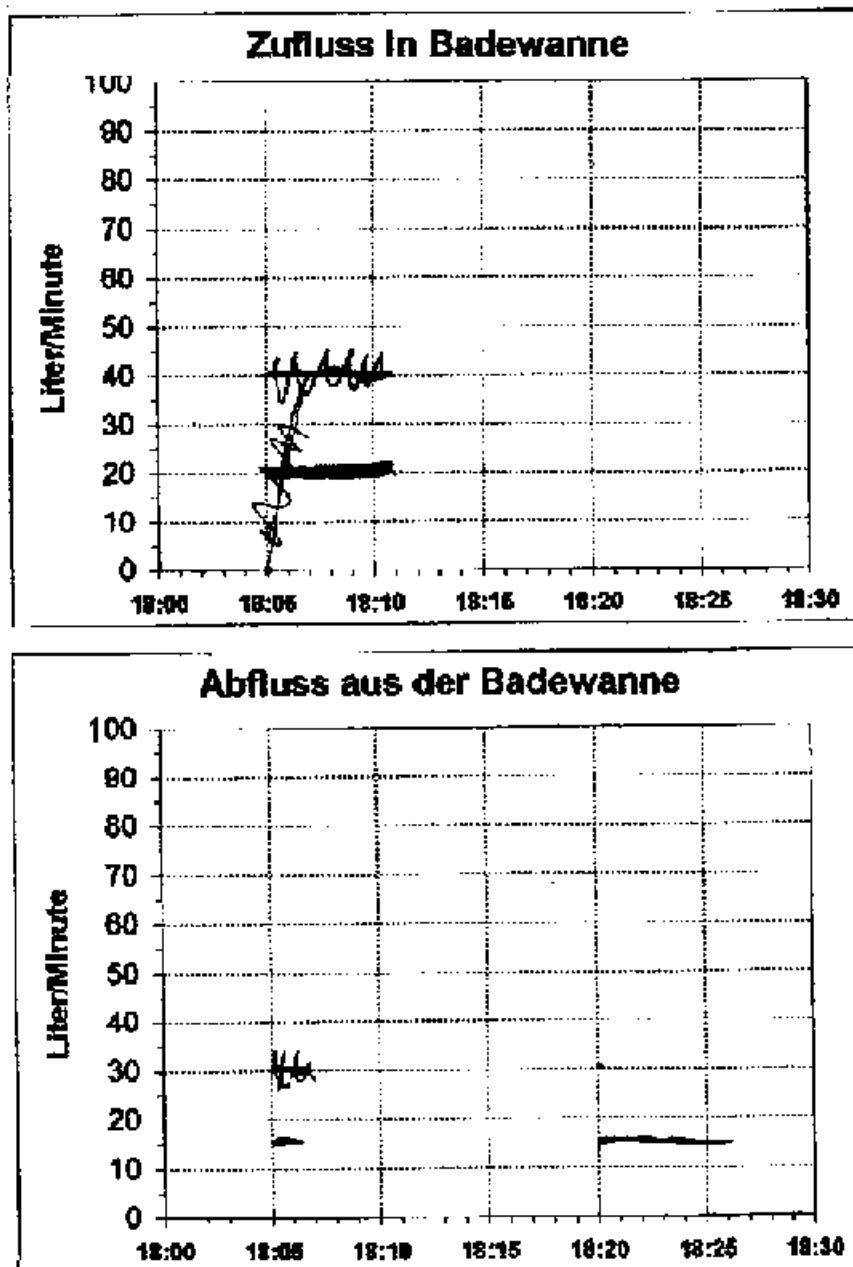


Abbildung 6-7: Aufgabe L - Lösung von "Data 6"

- 088 D Im oberen ist nur der Zufluss einzuzichnen und da er um 18 Uhr 5 den Wasserhahn aufdreht, beginnt eben ab 18 Uhr 5 20 Liter pro Minute einzurinnen – auf der y-Achse haben wir Liter pro Minute aufgetragen - und ich habe immer 20 Liter pro Minute – deswegen habe ich das horizontal eingezeichnet. Und um 18 Uhr 11 hat er den Wasserhahn zuge dreht und deswegen haben wir bis 18 Uhr 11 den Zufluss. Und unterhalb den Abfluss: Da haben wir von 18 Uhr 5 bis 18 Uhr 7 konstante 15 Liter, weil da die Pipe offen war, und von 18 Uhr 20 weg ist dann der Wasserhahn (Anm. I: gemeint war: Abfluss) wieder geöffnet und wir haben wieder 15 Liter pro Minute Abfluss)

7.3.3. Aufgabe F

Die zugehörige Aufgabe E wurde beim Vortest bereits von den meisten VPn richtig beantwortet. Diejenigen interviewten VPn mit den problematischen Lösungen beim Vortest waren genau die, die nicht mehr zum Nachinterview gekommen sind. Dementsprechend ist es wenig verwunderlich, dass bei Aufgabe F praktisch alle Lösungen beim Nachtest korrekt waren.

7.3.4. Aufgabe G

Beim Vortest haben eine Reihe von VPn zu Aufgabe B angegeben, dass das Maximum des Bestandes nicht ermittelt werden könne, weil nicht bekannt sei, wie viele Personen zu Beginn im Krankenhaus wären. VP "FM3" war eine davon. Beim Nachtest ermittelte Sie hingegen in Aufgabe G das Maximum des Autobestandes aus der Tabelle korrekt, indem Sie die ankommenden und abfahrenden Autos saldierte und damit rechnerisch ermittelte, dass um die Mittagszeit die meisten Autos am Parkplatz waren.

G) Die untenstehende Tabelle zeigt, wie viele Autos auf dem Uniparkplatz im Laufe eines Tages ankommen bzw. abfahren:

Stunde	An	Ab	Stunde	An	Ab
5-6	0	0	14-15	24	48
6-7	11	1	15-16	15	59
7-8	44	2	16-17	10	98
8-9	92	11	17-18	16	35
9-10	53	17	18-19	9	22
10-11	42	12	19-20	4	8
11-12	37	25	20-21	3	5
12-13	24	31	21-22	1	2
13-14	47	55	22-23	0	0

C-64

G1 Können Sie aus der Tabelle ermitteln, wann die meisten Autos auf dem Parkplatz stehen?

Ja, um 11-12 Uhr und zwar mit folgender Methode:

gr. An - A (gr. Ab)

Bestand am gr.

Nein, weil.....

(nach Abzug Ab v. Addition An)

G2 Wann sind die meisten Autos abgefahren?

zw. 16-17 Uhr

Abbildung 6-8: Aufgabe G - Lösung von "FM3"

7.3.5. Aufgabe H

Diese Aufgabe wurde beim Nachtest (anders als beim Vortest von den meisten Vpn korrekt gelöst. Vpn Gartner beispielsweise hat beim Vortest nur eine horizontale Linie bei Stand 100 gezeichnet. Beim anschließenden Interview erklärte Gartner die horizontale Linie damit, dass sich per Saldo die Zuflüsse und Abflüsse wieder ausgleichen und er nur mehr diesen Gesamt-Null-Effekt dann eingezeichnet habe.

Beim Nachtest hingegen hat Gartner den Bestand an Wasser in Beispiel H korrekt eingezeichnet und erklärt.

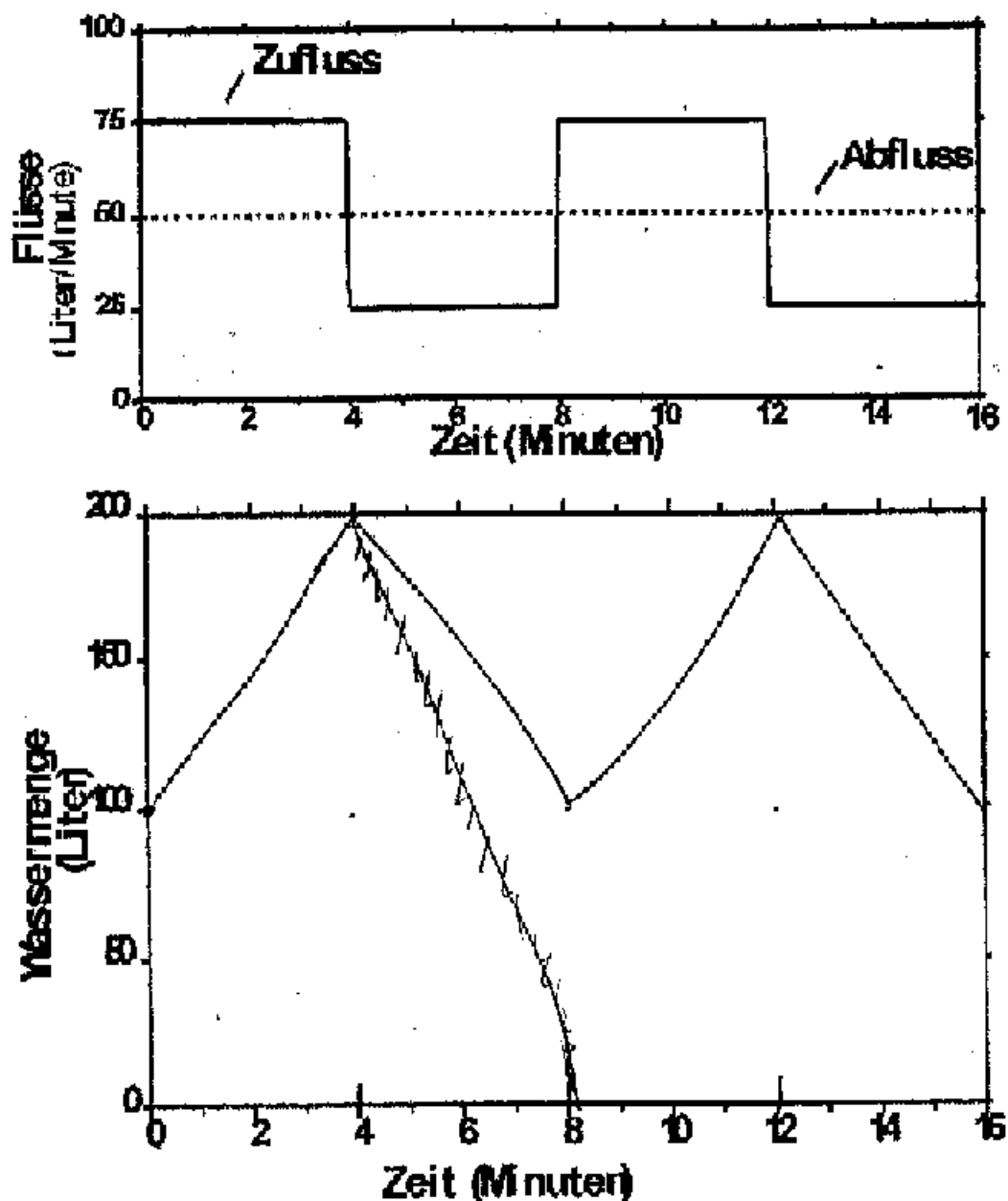


Abbildung 6-9: Aufgabe H - Lösung von "Gartner"

Dass nicht alle VPn beim Nachtest Aufgabe H problemlos bewältigten, belegt VPn FM3. Hier ist der Verlauf ab Minute 4 vom Zick-Zack-Muster her tendenziell richtig. Warum die VPn in den vier Minuten eine sinkende Wassermenge eingezeichnet hat, konnte sie beim anschließenden Interview nicht erklären, sondern meinte, sie habe hier einfach "irgend etwas" eingezeichnet, ohne dies jetzt nachträglich genauer begründen zu können.

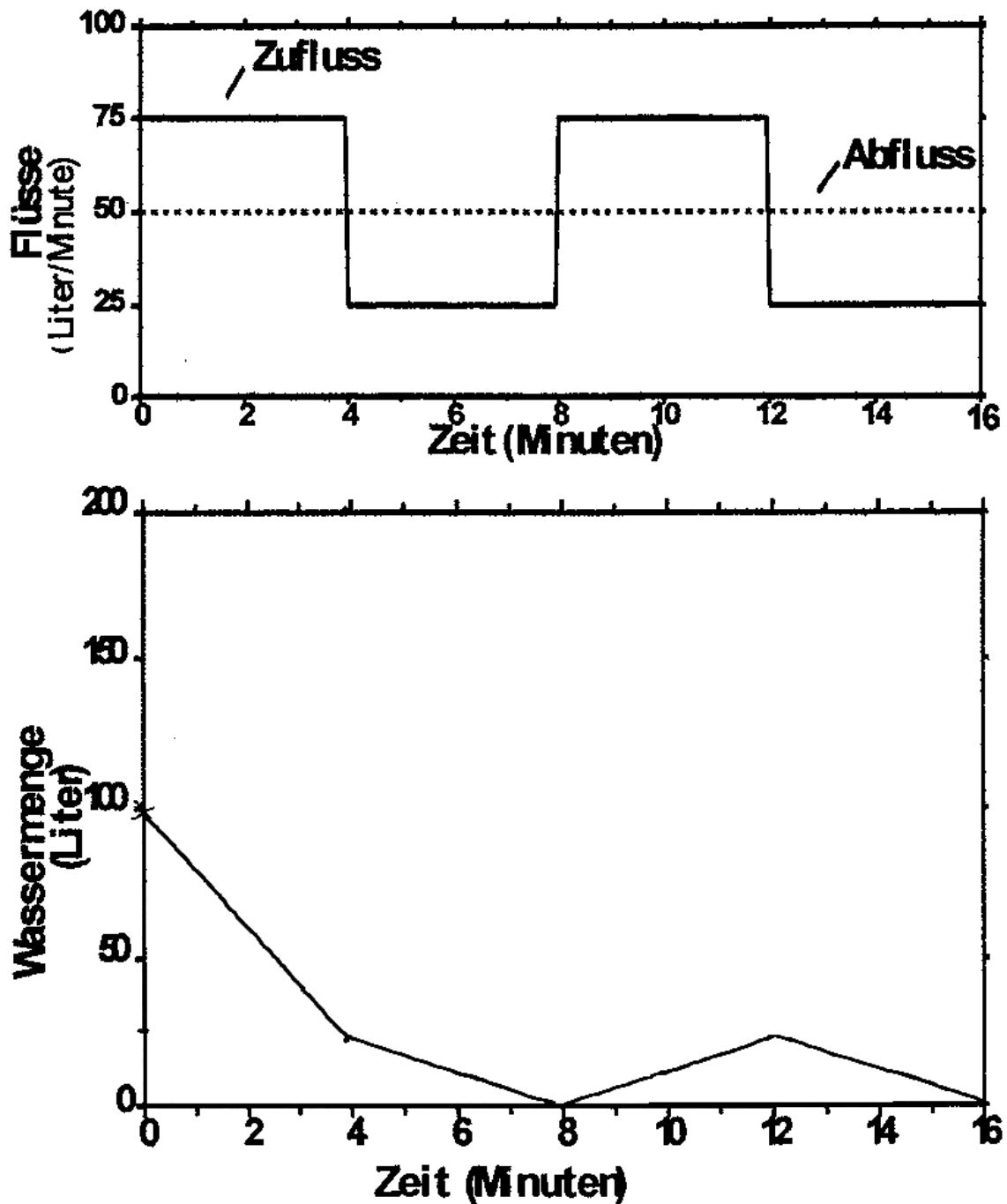


Abbildung 6-10: Aufgabe H - Lösung von "FM 3"

7.3.6. NachwortWie ist es uns als Interviewerinnen gegangen?

Für uns war dieses Projekt die erste Erfahrung mit empirischen Interviews. Es fiel uns zu Beginn nicht leicht, neutrale Formulierungen in den Fragen zu finden und keine Suggestivfragen zu stellen. Insgesamt haben wir es z.T. als spannend erlebt, die Hintergründe von verschiedenartigen Lösungen heraus zu bekommen. Eine etwas umfangreichere persönliche Stellungnahme erfolgt in einem separaten Bericht.

8. Resümee

(Ulrike Kotzent-Pietschnig)

Das Projekt "Unterscheidung von Bestands- und Flussgrößen" wurde von einem Team von Studierenden der Universität Klagenfurt unter der Leitung von Ao. Univ. Prof. Dr. Günther Ossimitz durchgeführt. Es ist damit die dritte größere Untersuchung zum Thema; nach der bahnbrechenden Pionierarbeit von Sweeney/Sterman (2000) und der ersten Studie im Deutschen Sprachraum von Ossimitz (2001a, 2001b). Erstmals wurde bei der vorliegenden Untersuchung ein Pretest-Posttest-Design gewählt, um die Auswirkungen eines ca. zweistündigen "Crash-Kurses" auf die Leistungen der Vpn zu untersuchen.

Warum ist der Untersuchungsgegenstand "Unterscheidung von Bestands- und Flussgrößen" von besonderem Interesse? Das Verwechseln von Beständen und Flüssen führt sehr leicht dazu, dass Daten falsch interpretiert werden. Dies kann in vielen Gebieten, wie z. B. Wirtschaft/Management, Biologie, Systemdynamik, Kybernetik, Sozialwissenschaften und Psychologie zu verheerenden Folgen in der Praxis führen. Eines der dramatischsten Beispiele dafür ist die irrige Meinung, das Erreichen eines Nulldefizites nach vielen Jahren permanenter Budgetdefizite gleichbedeutend mit dem Verschwinden des Schuldenbestandes ist. Versteht man systemisches Denken in einem umfassenden Sinn, so inkludiert dies auch die Förderung eines Verständnisses von zeitlich ablaufenden Prozessen, was die Fähigkeit zur Unterscheidung von Beständen und Flüssen zwingend bedingt. Das Interesse für systemisches Denken war schließlich auch für die studentischen Mitarbeiter(innen) eine wesentliche Motivation, sich im diesem Projekt zu engagieren.

8.1. Inwieweit wurden die Ziele des Projektes erreicht?

Ziel 1:

Es sollte ermittelt werden, inwieweit Studierende der Betriebswirtschaft der Universität Klagenfurt in einem Pretest Bestands- Flussgrößen unterscheiden verschiedenen Angaben und Aufgabenstellungen unterscheiden können.

Die beiden Vorgängeruntersuchungen von Sweeney/Sterman (2000) und in noch stärkerem Maße die Untersuchung von Ossimitz (2001a, 2001b) lieferten beide ernüchternde bis verheerende Resultate bei der Fähigkeit, Bestände und Flüsse zu unterscheiden. So darf es nicht verwundern, dass sich auch der Pretest im vorliegenden Projekt "Unterscheidung von Bestands- und Flussgrößen" nahtlos an die Erfahrungen der Vorgängerstudien anschließt. Auch unsere Studie zeigt ganz deutlich auf, eigentlich nur, dass die Unterscheidung zwischen Beständen und zugehörigen Zu- und Abflüssen vielen (überdurchschnittlich gebildeten!) Menschen große Schwierigkeiten bereitet und dies somit als ein gewichtiges Defizit in unseren "fortschrittlichen" Denkfähigkeiten bezeichnet werden kann. Für spezifischere Aussagen (z.B. geschlechts- oder altersspezifische Differenzierungen oder eine Differenzierung hinsichtlich der Vorbildung) war die Datenbasis mit n=94 VPn jedoch etwas zu knapp.

Ziel 2:

Nach dem Pretest sollte in einem "Crash-Kurs" von ca. 1 – 2 Stunden Dauer praktisch erprobt werden, inwieweit Basiskennnisse zur Unterscheidung von Bestands- und Flussgrößen praktisch vermittelt werden können.

Der Crash-Kurs war für die teilnehmenden Studierenden wirklich interessant. Er fand im Rahmen der im SS 2001 an der Uni Klagenfurt gehaltenen Vorlesung "Angewandte Statistik für Betriebswirte" statt, deren Teilnehmer auch gleichzeitig die untersuchte Personengruppe waren. Das Hauptproblem mit dem Crash-Kurs war, dass der knapp zweistündige Vorlesungsnachmittag mit dem Crash-Kurs nicht von allen Testteilnehmern besucht wurde. Dies wurde wenigstens teilweise dadurch kompensiert, dass es zum Crash-Kurs eine schriftliche Unterlage gab, die frei verfügbar war⁹.

Ziel 3:

In einem Nachtest (sog. Posttest) sollte festgestellt werden, in wie weit die Fähigkeit zur Unterscheidung von Bestands- und Flussgrößen sich durch den Crash-Test verbessert hatten.

⁹ <http://www.uni-klu.ac.at/~gossimit/pap/bestflus.pdf>

Zunächst kann festgestellt werden, dass die Leistungen beim Posttest zwar in einigen Teilbereichen deutlich besser als beim Pretest waren, dass man aber insgesamt auch beim Posttest noch nicht von einer wirklich überwältigend guten Leistung der VPn sprechen kann. Verschiedene Aspekte, die vornehmlich im keineswegs vollkommenen Untersuchungsdesign begründet liegen, erschweren klarerer Aussagen:

- Zunächst einmal ist die Zahl der untersuchten Personen oftmals zu gering, um auch bei bestehenden Zusammenhängen diese signifikant nachzuweisen.
- Fast ein Drittel der Vpn haben zwischen Pretest und Posttest aufgehört, so dass von diesen Leuten kein Nachtest mehr verfügbar war. Überdies zeigten die Leute, von denen wir keinen Nachtest mehr haben, bei einer Reihe von Fragen beim Vortest deutlich überdurchschnittliche Leistungen. Eine Beschränkung auf diejenigen Personen, die Vortest und Nachtest geschrieben haben, würde das Gesamtergebnis des Vortests drücken und damit einen größeren Leistungszuwachs (um den Preis einer schmaleren Beurteilungsbasis beim Pretest) zwischen Vor- und Nachtest ergeben.
- Eine gewisse Anzahl von Vpn haben den Nachtest absolviert, ohne den Crash-Kurs tatsächlich besucht zu haben.

8.2. *Interessante Ergebnisse*

Im Laufe dieses Projektes haben wir einige interessante Fakten gesammelt, die wir dann natürlich genauer untersuchen wollten, sei es durch ergänzende Interviews oder durch statistische Auswertungen mittels Chi-Quadrat-Unabhängigkeits-Tests. Hier nun eine kleine Zusammenfassung der wichtigsten Ergebnisse:

- Durchwegs war erkennbar, dass die jüngeren Jahrgänge (1978 – 82) bei dem Test besser abgeschnitten haben. Ob dies mit der Schulbildung zu tun hat, müsste erst weiter untersucht werden.
- Beide Tests beinhalteten eine 5-teilige Selbsteinschätzungs-Skala. Bei deren Auswertung festzustellen war, dass die männlichen Teilnehmer eher dazu neigen, ihre eigenen "skills" höher einzuschätzen, als es die weiblichen tun.

- Ob nun die Männer tatsächlich besser waren als die Frauen, konnte kaum schlüssig beantwortet werden. Bei einigen Items gab es leichte Tendenzen in diese Richtung, die jedoch aufgrund der sehr niedrigen Anzahl an untersuchten Personen kaum die Schwelle zu einem signifikanten Zusammenhang überschritt. Um diese Relation genauer zu untersuchen, bräuchten wir einen viel größeren Befragtenkreis.
- Die tatsächliche Qualität der Antworten hat mit Selbsteinschätzung der Teilnehmer im Lesen und Interpretieren von Graphiken gar nichts zu tun. Im Gegenteil: bei einigen Fragen schneiden sogar diejenigen Personen, die sich als schwach bis sehr schwach im Umgang mit Graphiken einschätzt haben, besser ab als diejenigen, die sich selbst als gute Interpretierer von Graphiken einschätzten.
- Weiters war zu erkennen, dass die Testpersonen bei Tabellendarstellungen von Zu- und Abflüssen viel öfter zu einer korrekten Interpretation des daraus resultierenden Bestandes kamen, als bei Graphikdarstellungen. Dies liegt wohl daran, dass man bei einer tabellarischen Angabe mechanisches Saldieren (Addieren der Zuflüsse, Subtrahieren der Abflüsse) die gestellten Aufgaben lösen konnte, auch wenn der Unterschied zwischen einer Bestands- und Flussgröße der betreffenden Vpn gar nicht bewusst war. Somit können wir sagen, dass Graphikdarstellungen hinsichtlich der Gefahr, Flüsse irrtümlich als Bestände zu interpretieren, als problematischer einzustufen sind.

9. Ausblick

(Ulrike Kotzent-Pietschnig)

In diesem Abschnitt unseres Projektberichtes möchten wir nun diejenigen Dinge behandeln, die wir unserer Meinung nach noch gerne genauer untersucht hätten, oder die wir für verbesserungswürdig halten.

- Weiter untersucht hätte wir sehr gerne den **schulischen Background** der Versuchspersonen. Da wir glauben, dass es bestimmte Unterschiede zwischen den vorher besuchten Schultypen (BG, HAK, HBLA, usw.) und auch der Besuchszeit gibt (ältere bzw. jüngere Jahrgänge). Damit wir der Frage nachgehen können, ob die Schulausbildung (z. B. in bestimmten Fächern) früher vielleicht besser war als heutzutage. Oder ob sich die persönliche Einstellung zum Lernen geändert hat.
- Auch interessant wäre eine **Ausweitung der Untersuchung** auf noch mehr Versuchspersonen gewesen. Durch eine Vergrößerung des Befragtenkreises könnte man auch genauer auf die Frage des Zusammenhangs zwischen Antwortverhalten und Geschlecht eingehen. Bei dem geringen Stichprobenumfang lassen sich kaum signifikante Zusammenhänge nachweisen, aber wir würden gerne wissen, zu welchen Ergebnissen dieselben Hypothesen bei einer größeren VP-Anzahl führen.
- Für das nächste Projekt würden wir ein **professionelleres Design über Bezahlung** vorschlagen, um damit die Probleme mit Ausfällen oder Nicht-Anwesenheit beim Crash-Kurs zu minimieren. Eigentlich ist nämlich der Vergleich zwischen Pre- und Posttest nur für jene Personen relevant, die beide Tests mitgeschrieben haben. Denn nur so könnte man eine wirkliche Verbesserung (oder vielleicht auch Verschlechterung) nach dem Crash-Kurs feststellen.
- Die **persönliche Einschätzung** nach dem Crash-Kurs wäre auch noch ein interessanter Aspekt für unsere Untersuchung gewesen. Wie sich die Personen nach dem Crash-Kurs persönlich eingeschätzt hätten. Dies hätte man in neuerlichen Selbsteinschätzungsfragen im Posttest feststellen können.

Die Qualität eines Projektes ist großteils davon abhängig, wie gut es durchorganisiert wird, wie gut die Zeitpläne sind und wie genau diese eingehalten werden. Es gibt viele mögliche Projekt-Layouts, doch sobald man bestimmte Daten gesammelt hat, hat man auch immer das Problem der Auswertung und der Verwertung von den gesammelten Informationen. Wir sind der Meinung, dass wir diese Fragen innerhalb unseres Projektteams mit transparenten und leicht verständlichen Bewertungsrichtlinien sehr gut managen konnten. Weiters muss man aber auch sagen, dass die Zusammenarbeit zwischen den einzelnen Projektmitarbeitern tadellos funktionierte und man zu keinem Zeitpunkt das Gefühl hatte, dass unsere Mitarbeiter nicht motiviert waren. So konnten wir auch die Bewertung des Posttestes in kürzester Zeit durchziehen und zu Semesterende eine komplette Computerpräsentation der ersten Ergebnisse zeitgerecht vorstellen.

Wir hoffen nun, dass wir unseren Lesern das Thema bzw. die Unterscheidung zwischen "Fluss- und Bestandsgrößen" etwas näher bringen konnten. Und dies in Zukunft hilft einige Komplikationen oder Verwechslungen mit Beständen und Flüssen zu verringern. In diesem Sinne : "AUSGELERNT HAT MAN NIE!!"

10. Nachwort des Projektleiters

(Günther Ossimitz)

Das vorliegende Projekt "Unterscheidung von Bestands- und Flussgrößen" war in mehrerer Hinsicht bemerkenswert. Auf der Ebene empirischer Forschung zum Thema "Stock-Flow-Thinking" wurde mit dem dreistufigen Design Pretest – Crash-Kurs – Posttest völliges Neuland betreten. Es zeigte sich, dass seriöse empirische Forschung in dem (an sich schon recht anspruchsvollem) Feld "Stock-Flow-Thinking" vor ganz gewaltigen forschungsmethodologischen Herausforderungen steht, die in der vorliegenden Studie zum ersten Mal praktisch erlebt, aber sicherlich noch nicht wirklich befriedigend bewältigt werden konnten.

Auf der Ebene der Projektdurchführung habe ich erstmals systematisch ein Projekt mit einer großen Zahl freiwilliger studentischer Projektmitarbeiter(innen) durchgezogen. In der Tat ist der große Umfang dieses Projekts ganz wesentlich darauf zurückzuführen, dass sich im März 2001 bei einer Vorbesprechung über 15 Studierende sich darum rissen, an dem angekündigten Projekt ehrenamtlich mitarbeiten zu dürfen. Diese Studierenden haben sich mit z.T. tollkühnen Mut und bemerkenswertem Eifer in eine für sie vollkommen neue Aufgabe hineinbegeben, die sie (als Studierende jüngerer Semester!) bis an die Grenze ihrer Leistungsfähigkeit forderte. Mit großer Freude darf ich feststellen, dass es entgegen aller sonstiger Erfahrungen mit dem Durchhaltewillen von Studierenden hier bis zum Schluss trotz härtester Belastungen keinen einzigen Ausfall gegeben hat!

Ich habe als Projektleiter versucht, die Dynamik und Initiative der Studierenden nicht einzuengen, sondern sie nur behutsam in möglichst geordnete Bahnen zu lenken. Dies impliziert einen sehr hohen Anteil an Eigeninitiative in der Gestaltung der einzelnen Projektschritte und spiegelt sich auch in einer z.T. recht persönlichen Note in manchen Passagen des Projektberichts wider. Ich sehe diesen für einen Projektbericht vielleicht etwas ungewöhnlichen Touch als etwas Positives an, weil er authentisch ist und auch den besonderen Charakter dieses Projekts widerspiegelt.

Einen wirklich hervorragenden Job haben Benjamin Kreisler und Ulrike Kotzent-Pietschnig – mit kräftiger Unterstützung von Martin Waiguny und Melanie Zoltan - als "Stabsstelle" bei der Koordination der vielfältigen Projektaktivitäten der einzelnen Mitarbeiter(innen) geleistet. Dieses Quartett hat hier nicht nur ein höchst erfolgreiches Praktikum im Projekt-Management absolviert, sondern dann auch noch unter verschärften Bedingungen (räumliche Trennung in den Ferien) tüchtig den vorliegenden Projektbericht erstellt. Ergänzt wurde dieses "professionelle" Projektteam noch durch die Damen Sandra Galli, Eva Gratzler und Renate Weichlinger, die mit einer für Newcomer wirklich bemerkenswerten Courage und voller Elan die ergänzenden Interviews bzw. die Gestaltung der Web-Page zum Projekt managten.

Dennoch dürfen an dieser Stelle keineswegs die zehn "ehrenamtlichen" Mitarbeiter(innen) Josef Biedermann, Ingrid Dogan, Edzo Frerichs, Andrea Grabner, Peter Ladinig, Reiner Micheler, Christian Petautschnig, Michael Rachbauer, Markus Steinbrugger und Philipp Tillian vergessen werden. Sie haben in Zweierteams je eine Frage des Vor- und Nachtests ausgewertet und mit ihrer Dokumentation die Arbeit des engeren Projektteams entscheidend unterstützt.

Insgesamt bleibt mir nur zu hoffen, dass das Projekt „Bestands- und Flussgrößen“ sowohl für die studentischen Mitarbeiter(innen) als auch für die rezipierende wissenschaftliche Community genauso attraktiv und spannend erlebt wird wie von mir.

Literaturverzeichnis

Dörner, Dietrich (1989): *Die Logik des Misslingens*. Reinbek: Rowohlt.

Dörner, Dietrich / Franz Reither (1978): *Über das Problemlösen in sehr komplexen Realitätsbereichen*. Zeitschrift für Experimentelle und Angewandte Psychologie, 25, 527-551.

Funke, Joachim (1984): *Diagnose der westdeutschen Problemlöseforschung in Form einiger Thesen*. Sprache und Kognition, 3, 1984(3), 159-172.

Funke, Joachim (1985): *Problemlösen in komplexen computersimulierten Realitätsbereichen*. Sprache und Kognition, 1985(3), 113-129.

Gomez, Peter / Gilbert Probst (1987): *Vernetztes Denken im Management: Eine Methodik des ganzheitlichen Problemlösens* (Die Orientierung 89). Bern: Schweizerische Volksbank.

Gomez, Peter / Gilbert Probst (1997): *Die Praxis des ganzheitlichen Problemlösens: Vernetzt denken, Unternehmerisch handeln, Persönlich überzeugen* (2 Aufl.). Bern: Paul Haupt.

Ossimitz, Günther (1995): *Entwicklung vernetzten Denkens bei Schülern der S II*. In: K. P. Müller (Hg.): Beiträge zum Mathematikunterricht (pp. 364-367).

Ossimitz, Günther (1996): *Stand und Perspektiven der Forschung zum systemischen Denken*. In: G. Kadunz et al. (Hg.): Trends und Perspektiven (Bd. 23, pp. 279-286). Wien: Hölder-Pichler-Tempsky.

Ossimitz, Günther (2000): *Entwicklung systemischen Denkens* München: Profil Verlag.

Ossimitz, Günther (2001a): *Unterscheidung von Bestands- und Bewegungsmassen: Ergebnisse einer empirischen Untersuchung*. In: Gabriele Kaiser (Hg.): Beiträge zum Mathematikunterricht 2001 (pp. 468-472).

Ossimitz, Günther (2001b): *Stock-Flow-Thinking and Reading stock-flow-related Graphs: An Empirical Investigation in Dynamic Thinking Abilities*. Manuskript, Universität Klagenfurt. <http://www.uni-klu.ac.at/~gossimit/pap/sfthink.pdf>

Probst, Gilbert / Peter Gomez (Hg.) (1991): *Vernetztes Denken: Ganzheitliches Führen in der Praxis* (2 Aufl.) Wiesbaden: Gabler Verlag.

Richmond, Barry (1991): *Systems Thinking: Four Key Questions*. Lyme, NH: High Performance Systems Inc.

URL: <ftp://sysdyn.mit.edu/ftp/cle/documents/system-ed/SE1993-05STFourKeyQuestions.pdf>

Richmond, Barry (1993): *Systems thinking: critical thinking skills for the 1990s and beyond*. System Dynamics Review, 9(2), 113-133.

URL: <ftp://sysdyn.mit.edu/ftp/cle/documents/system-ed/SE1993-05STCriticalThinking.pdf>

Richmond, Barry (1994): *System Dynamics/Systems Thinking: Let's just Get On With It*. Lyme, NH: High Performance Systems, Inc.

Senge, Peter M. (1990): *The Fifth Discipline: the Art and Practice of the Learning Organization* (1st Aufl.). New York: Doubleday.

Sweeney, Linda Booth / John D. Sterman (2000): *Bathtub Dynamics: Initial Results of a Systems Thinking Inventory.* System Dynamics Review, (16:4), pp 249 - 286.

Vester, Frederic (1983): *Unsere Welt - ein vernetztes System.* München: Deutscher Taschenbuch Verlag.

Vester, Frederic (1984): *Neuland des Denkens.* München: Deutscher Taschenbuch Verlag.

Vester, Frederic (1988): *Leitmotiv vernetztes Denken.* München: Heyne Verlag.

Vester, Frederic (1999): *Die Kunst vernetzt zu denken. Ideen und Werkzeuge für einen neuen Umgang mit Komplexität.* Stuttgart: DVA.