

- 3 In Anlehnung an PISA, setzen wir als Kriterium für die Beherrschung einer Testaufgabe eine Lösungswahrscheinlichkeit von 65 Prozent an. Die in Tabelle 3.1 eingetragenen Schwierigkeitskennwerte bezeichnen jenen Punkt auf der Kompetenzskala, an dem diese Lösungswahrscheinlichkeit erreicht wird.

Joachim Wirth und Joachim Funke

4 Dynamisches Problemlösen: Entwicklung und Evaluation eines neuen Messverfahrens zum Steuern komplexer Systeme

4.1 Einleitung

In Kapitel 1 wurden analytische und dynamische Probleme voneinander unterschieden. Bei analytischen Problemen sind zu Beginn alle für die Lösung relevanten Informationen über den aktuellen und den herbeizuführenden Zustand entweder gegeben, oder sie können auf der Basis der gegebenen Informationen erschlossen werden. Diese Informationen verändern sich im Laufe des Problemlöseprozesses nicht. Es sind *statische* Probleme, deren Lösung die Entscheidung für eine richtige Operation oder die korrekte Kombination mehrerer Operationen erfordert. Diese Entscheidung erfolgt auf der Grundlage genauer Analysen der Informationen, durch die der Anfangs- und der Zielzustand und die möglichen Zwischenzustände definiert sind. Die für diese Analysen notwendigen Aspekte des Problemlösens wurden bereits in Kapitel 3 ausführlich beschrieben.

Der andere Problemtyp ist durch die Dynamik und Interaktivität der Situation gekennzeichnet. Die Aufgabe beim dynamischen Problemlösen besteht in der Steuerung eines unbekanntes dynamischen und mehr oder weniger komplexen Systems. Zu Beginn sind die für die Systemsteuerung relevanten Informationen zu einem großen Teil weder verfügbar noch erschließbar, sondern müssen durch die Interaktion mit dem System erst in Form von Veränderungen des Systemzustands erzeugt werden. Der Problemlöser muss diese Informationen dann kontinuierlich verarbeiten und bewerten, ob die getätigten Eingriffe zum Ziel führen oder nicht. Da dynamische Systeme meist auch sehr komplex sind und das Wissen über ihre Struktur die Grenzen des Arbeitsgedächtnisses überschreitet, sind bei diesem Problemtyp Aspekte des (selbstregulierten) Lernens meist eine notwendige Voraussetzung für die erfolgreiche Steuerung des Systems.

Damit müssen beim dynamischen Problemlösen zwei Hauptanforderungen unterschieden werden, die durch unterschiedliche Prozesse und Handlungen zu bewältigen sind (vgl. Funke, 2003, Abschnitt 5.2). Zum einen muss Wissen über das System interaktiv erworben werden. Zum anderen muss Wissen im Rahmen der zielgerichteten Sys-

temsteuerung angewandt werden. In realen Problemsituationen müssen diese beiden Anforderungen meist gleichzeitig bewältigt werden. So kann zum Beispiel ein Bürgermeister nicht zunächst lange Zeit ausschließlich damit verbringen, die Strukturen seiner Stadt und aller ihrer Subsysteme kennen zu lernen, bevor er Entscheidungen trifft und beginnt, die (Sub-)Systeme zu regulieren (Dörner u.a., 1983; siehe auch Kap. 1). Er muss von Beginn an versuchen, die einzelnen Systeme in einen optimalen Zustand zu steuern und sie dort zu halten. Für eine genaue Diagnose der beiden Fähigkeitskomponenten des Wissenserwerbs und der Wissensanwendung empfiehlt es sich jedoch, sie getrennt voneinander zu messen.

Für die Erfassung dynamischer Aspekte von Problemlösefähigkeit besteht damit die Notwendigkeit, ein Verfahren zur Verfügung zu stellen, bei dem der Problemlöser zum einen seine Fähigkeit unter Beweis stellen kann, durch Interagieren mit einem unbekanntem komplexen und dynamischen System Wissen über dessen Struktur und die Möglichkeiten seiner Steuerung zu erwerben. Zum anderen muss er zeigen können, dass er das einmal erworbene Wissen auch für eine zielgerichtete Steuerung des Systems erfolgreich anzuwenden weiß. Für die Erfassung sowohl des interaktiven Wissenserwerbs als auch der zielgerichteten Wissensanwendung muss das Verfahren auf die jeweiligen Handlungen und Eingriffe des Problemlösers reagieren, indem es seinen Zustand ändert und dem Problemlöser Informationen über diese Zustandsänderungen zurückmeldet. Papier- und Bleistift-Verfahren, wie sie für die Erfassung analytischer Aspekte des Problemlösens genutzt werden konnten, sind nicht dazu in der Lage, diese Interaktionen zwischen Problemlöser und zu steuerndem System zu simulieren und zu erfassen. Es muss in diesem Fall auf computerbasierte Simulationen komplexer und dynamischer Systeme zurückgegriffen werden. Diese können so programmiert werden, dass sie genau diese Anforderungen erfüllen. Zudem bieten sie die Möglichkeit, nicht nur Produktmaße für den Erfolg des Wissenserwerbs und die Wissensanwendung zur Verfügung zu stellen. Zusätzlich können verhaltensbasierte Prozessmaße der Regulationsgüte beim Wissenserwerb und bei der Wissensanwendung implementiert werden. Solche Prozessmaße sind im Rahmen von *Large-Scale Assessments* wie PISA ein Novum und auch in der experimentellen psychologischen Forschung nur selten realisiert worden. Sie sind jedoch für das Aufdecken von Ansatzpunkten zur Förderung dynamischer Aspekte des Problemlösens von unschätzbarem Wert.

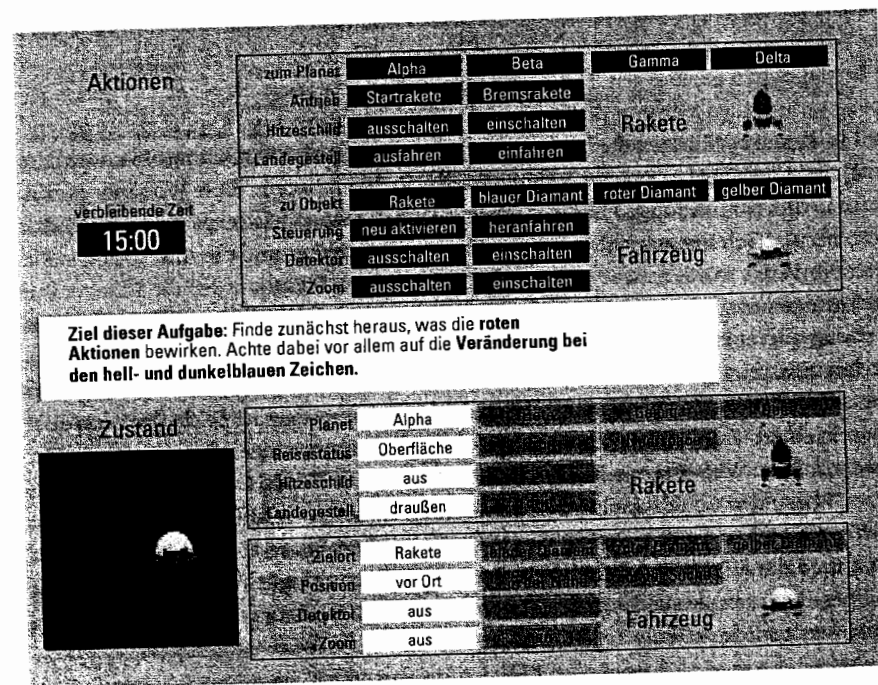
2 Instrument und Leistungsmaße

2.1 Komplexes und dynamisches System

Im Feldtest 1999 wurden insgesamt drei verschiedene computerbasierte Verfahren zur Erfassung dynamischer Problemlöseaspekte erprobt (siehe Kap. 2), von denen zwei in der Haupterhebung 2000 in weiter optimierten Versionen zum Einsatz kamen. Das war zum einen das Raumfahrtspiel (Funke, Töpfer, & Wagener, 1998), zum anderen das ur-

sprünglich von Schrettenbrunner entwickelte System „Hunger in Nordafrika“ (Leutner & Schrettenbrunner, 1989), das von Leutner für PISA adaptiert wurde (vgl. die Ausführungen zum Ökologie-Planspiel in Kap. 2). Die Struktur des Raumfahrtspiels ist formal vollständig beschreibbar, was für die Struktur des Ökologie-Planspiels aufgrund seiner hohen Komplexität nur schwer bzw. gar nicht möglich ist. Die vollständige formale Beschreibbarkeit des Raumfahrtspiels ermöglichte es, innovative Maße der Handlungsregulation zu konstruieren und somit tiefgehende und zuverlässige Analysen von Wissenserwerbs- und Wissensanwendungsprozessen durchzuführen. Diese Maße unterscheiden sich von den Maßen, wie sie im Feldtest 1999 erprobt wurden, und stellen eine konsequente Weiterentwicklung des dort erprobten Instruments dar. Kapitel 4 beschäftigt sich ausschließlich mit diesen weiterentwickelten Maßen. Auf die Darstellung der Ergebnisse, die mit dem System des Ökologie-Planspiels erzielt wurden, wird deshalb an dieser Stelle verzichtet.

Abbildung 4.1 zeigt die Bildschirmoberfläche des Raumfahrtspiels. Als zu steuernde Subsysteme stehen eine Rakete und ein Planetenfahrzeug zur Verfügung, die analoge Strukturen aufweisen. Mit der Rakete kann zwischen vier Planeten hin- und hergeflogen



Die hier schwarz abgebildeten Aktionen wurden in der Studie als rote Schaltflächen präsentiert. Die hier weiß abgebildeten Zeichen wurden in der Studie hellblau, die hier grauen Zeichen wurden in der Studie dunkelblau präsentiert.

Abbildung 4.1 Bildschirmoberfläche des Heidelberger Finiten Automaten

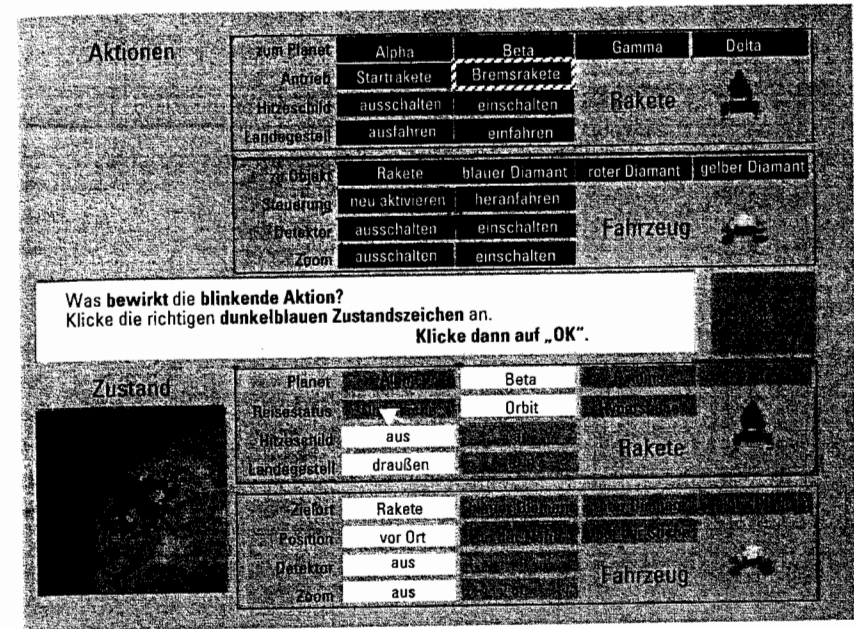
werden, auf jedem dieser Planeten kann mit dem Planetenfahrzeug zu jeweils drei Diamanten und der Rakete gefahren werden. In der oberen Hälfte des Bildschirms sind die Menge der Schalter (rote Schaltflächen) dargestellt. Die Rakete und das Fahrzeug werden gesteuert, indem mit der Computermaus diese roten Schalter angeklickt werden. In der unteren Hälfte geben blaue Anzeigen durch einen Farbwechsel von Dunkel- zu Hellblau an, in welchem Zustand sich die Rakete und das Planetenfahrzeug befinden. Diese Anzeigen ändern ihre Farbe nur als Reaktion auf das Drücken eines roten Schalters. Das direkte Anklicken dieser Ausgabesignale führt zu keiner Zustandsänderung des Systems. Im Beispiel in Abbildung 4.1 befindet sich die Rakete auf der Oberfläche des Planeten Alpha, ihr Hitzeschild ist ausgeschaltet und ihr Landegestell ist ausgefahren. Das Planetenfahrzeug befindet sich direkt bei der Rakete mit ausgeschaltetem Detektor und Zoom.

Das Raumfahrtspiel ist so konstruiert, dass es insgesamt 116 verschiedene Zustände annehmen kann, in denen jeweils 20 verschiedene Schalter als Eingriffsmöglichkeiten zur Verfügung stehen. Mit diesen insgesamt $116 \times 20 = 2.320$ verschiedenen Eingriffsmöglichkeiten simuliert das Raumfahrtspiel ein durchaus komplexes und dynamisches System.

Um die Fähigkeiten zum interaktiven Wissenserwerb einerseits und zur zielgerichteten Wissensanwendung andererseits separat voneinander messen zu können, hatten die Schülerinnen und Schüler zuerst 15 Minuten Zeit, das Raumfahrtspiel frei und selbstständig zu explorieren. Die in der Haupterhebung 2000 eingesetzte Version des Spiels unterschied sich von der Version des Feldtests 1999 (vgl. Kap. 2) dahingehend, dass in diesen 15 Minuten die beiden Subsysteme des Spiels (Rakete bzw. Planetenfahrzeug) nicht durchgehend zu explorieren waren. Vielmehr wurde diese Wissenserwerbsphase in drei fünfminütige Phasen aufgeteilt. In der ersten Phase explorierten die Schülerinnen und Schüler ausschließlich die Rakete; die Schalter und die Anzeigen des Planetenfahrzeugs wurden nicht präsentiert. In der zweiten Phase wurden die Schalter und die Anzeigen der Rakete ausgeblendet, und es konnte ausschließlich mit dem Planetenfahrzeug interagiert werden. In der dritten Phase standen schließlich beide Subsysteme für den interaktiven Wissenserwerb zur Verfügung. Durch diese Aufteilung wurde zum einen der selbstregulierte interaktive Wissenserwerb gegenüber der Feldtestversion deutlich erleichtert. Zum anderen konnte so der Einfluss unterschiedlicher Arten von Vorwissen auf den Wissenserwerb untersucht werden (siehe Wirth, 2004): Während in der ersten Phase den Schülerinnen und Schülern kein systemspezifisches Vorwissen zur Verfügung stehen konnte, verfügten sie in der zweiten Phase aufgrund der analogen Strukturen der Rakete und des Planetenfahrzeugs über analoges Vorwissen. In der dritten Phase konnten sie auf das in den ersten beiden Phasen erworbene systemspezifische Vorwissen zurückgreifen.

4.2.2 Erfolg des Wissenserwerbs und der Wissensanwendung

Welches und wie viel Wissen sich die Schülerinnen und Schüler in den drei Phasen des Wissenserwerbs aneignen konnten, wurde direkt im Anschluss an die dritte Phase des



Die hier schwarz abgebildeten Aktionen wurden in der Studie als rote Schaltflächen präsentiert. Die hier weiß abgebildeten Zeichen wurden in der Studie hellblau, die hier grauen Zeichen wurden in der Studie dunkelblau präsentiert. Bei der Aufgabenpräsentation blinkte der gestreift umrandete Schalter.

Abbildung 4.2a Aufgaben der Erfolgsmaße, prospektive Aufgabe

Wissenserwerbs in einem Wissenserwerbtest gemessen. Der Test bestand aus jeweils acht Interpolations- und prospektiven Aufgaben (Funke & Buchner, 1992; siehe Abb. 4.2a). Danach wurden die Jugendlichen vor bis zu 22 Steuerungsaufgaben gestellt, in denen sie jeweils ihr Wissen zur Erreichung des Aufgabenziels anwenden mussten (Abb. 4.2b). Der prozentuale Anteil richtig beantworteter Interpolations- und prospektiver Aufgaben ist ein *Erfolgsmaß für den interaktiven Wissenserwerb*, der Anteil erreichter Steuerungsziele ein *Erfolgsmaß für die zielgerichtete Anwendung* des erworbenen Wissens. Beide Indikatoren für die Fähigkeit zum interaktiven Wissenserwerb bzw. für die Fähigkeit zur zielgerichteten Wissensanwendung wurden bereits im Feldtest erprobt und hatten sich dort bewährt (vgl. Kap. 2).

Erfolgsmaße messen ausschließlich den Erfolg. Sie sind nicht geeignet, den Weg zu beschreiben, der zu diesem Erfolg geführt hat. Will man jedoch nicht nur Aussagen darüber treffen, ob eine Person erfolgreich war oder nicht, sondern auch Erkenntnisse darüber gewinnen, wieso sie erfolgreich oder nicht erfolgreich war, sind Informationen über diesen Weg und seine Güte notwendig. Diese können zum Beispiel durch den Einsatz von Regulationsmaßen gewonnen werden.

Aktionen		zum Planet	Alpha	Beta	Gamma	Delta
Antrieb		Startrakete		Bremsrakete		
Hirzschild		ausschalten		einschalten		Rakete
Landegestell		ausfahren		einfahren		
		zu Objekt	Rakete	blauer Diamant	roter Diamant	gelber Diamant
Steuerung		neu aktivieren		heranfahen		
Detektor		ausschalten		einschalten		Fahrzeug
Zoom		ausschalten		einschalten		

Bringe die Rakete und das Fahrzeug zum blinkenden Ziel. Dazu benötigst Du jetzt mehrere Aktionen. Wenn Du nicht weiter kommst, kannst Du mit dem grünen Knopf eine neue Aufgabe beginnen.

Zustand		Planet	Alpha	Beta	Gamma	Delta
Reisestatus		Oberfläche				
Hirzschild		aus				
Landegestell		draußen				Rakete
		Zielort	Rakete	blauer Diamant	roter Diamant	gelber Diamant
Position		vor Dir		in der Nähe		
Detektor		aus				Fahrzeug
Zoom		aus				

Die hier schwarz abgebildeten Aktionen wurden in der Studie als rote Schaltflächen präsentiert. Die hier weiß abgebildeten Zeichen wurden in der Studie hellblau, die hier grauen Zeichen wurden in der Studie dunkelblau präsentiert. Bei der Aufgabenpräsentation blinkten die schwarz umrandeten Anzeigen.

Abbildung 4.2b Aufgaben der Erfolgsmaße, Steuerungsaufgabe

2.3 Regulation des Wissenserwerbs und der Wissensanwendung

Für den Prozess des Wissenserwerbs wurde ein Maß entwickelt (so genanntes „log_(or)“-Maß; gesprochen: „log-o-r-Maß“), das auf der Basis von Verhaltensdaten die selbstbestimmte Lernregulation erfasst (Wirth, 2004; siehe auch Anhang A). Das Maß basiert auf der Unterscheidung zweier Anforderungen, die ein Lerner beim interaktiven Wissenserwerb zu bewältigen hat. Zum einen muss er neu zu erwerbende Informationen im Umgang mit dem System entdecken bzw. in Form von Feedbackinformationen generieren. Die dafür notwendigen Lernhandlungen werden unter dem Begriff „Identifizieren“ gefasst. Strategien und Vorgehensweisen, die erfolgreich mit diesem Ziel eingesetzt werden können, werden in der Literatur zum *Scientific Discovery Learning* (z.B. Klahr & Dunbar, 1988; van Joolingen & de Jong, 1997) bzw. des *Inquiry Learning* (z.B. Keselman, 2003; White, Shimoda, & Frederiksen, 2000) beschrieben. Zum anderen muss der Lerner dafür Sorge tragen, dass er einmal identifizierte Informationen nicht wieder vergisst, sondern in späteren Anwendungssituationen sicher und leicht abrufen und nutzen kann. Lerntätigkeiten, die mit diesem Ziel ausgeübt werden, können unter

dem Begriff „Integrieren“ zusammengefasst werden. Strategien hierfür werden in der Literatur zum selbstregulierten Lernen (z.B. Artelt, 2000; Baumert & Köller, 1996; Mandl & Friedrich, 1992; Schreiber, 1998) bzw. zum Erwerb von Fertigkeiten (z.B. Anderson, 1983, 1996) untersucht. Lernhandlungen, die dem Identifizieren neuer Informationen dienen, unterscheiden sich jedoch von Lernhandlungen, die das Integrieren zum Ziel haben. Insofern steht ein Lerner im Umgang mit einem unbekanntem komplexen und dynamischen System fortwährend vor der Entscheidung, *entweder* identifizierende *oder* integrierende Eingriffe auszuüben, wodurch sich die Notwendigkeit zur Handlungsregulation ergibt.

Das log_(or)-Maß erfasst für bestimmte Zeitintervalle, ob ein Lerner den Lernprozess stärker auf das Identifizieren von Informationen ausrichtet oder stärker das Integrieren von Informationen zum Ziel hat. Wird der gesamte Lernprozess in mehrere Zeitintervalle unterteilt, kann damit der Verlauf der Lernprozessregulation beschrieben werden. Für alle drei Phasen des Wissenserwerbs wurde jeweils ein solcher Regulationsverlauf abgebildet und mithilfe latenter Wachstumskurven modelliert (Wirth, 2004; siehe auch Anhang A). In diesem Bericht beschränken wir uns auf die Darstellung des Regulationsverlaufs der mittleren Phase, in der analoges Vorwissen zur Verfügung stand. In dieser Phase kann die Lernprozessregulation am zuverlässigsten erhoben werden, da für diese Phase keine unkontrollierten Vorwissenseffekte angenommen werden müssen. Die Analysen zu Vorwissenseffekten sowie detaillierte Darstellungen der Regulationsverläufe in der ersten und dritten Phase sind an anderer Stelle bereits veröffentlicht (Wirth, 2004). Als Gütemaß für die Regulation des Wissenserwerbs wird der Wert des Steigungsfaktors des latenten Wachstumskurvenmodells für die mittlere Phase herangezogen (Anhang A; siehe auch Abschnitt 4.5). Dieser Steigungsfaktor ist ein echtes verhaltensbasiertes Prozessmaß und damit ein Novum im Rahmen von *Large-Scale Assessments*.

Eine optimale Regulation der zielgerichteten Wissensanwendung ist dann gegeben, wenn ein Schüler oder eine Schülerin sein bzw. ihr erlerntes Wissen über Eingriffsmöglichkeiten und ihre Effekte immer so nutzt, dass durch einen Eingriff der Abstand zum vorgegebenen Ziel kleiner wird. Für die Bewertung der Regulationsgüte bei der Wissensanwendung wurde daher ein dem log_(or)-Maß für den Wissenserwerb entsprechendes log_(or)-Maß für die Wissensanwendung entwickelt, das einen Wert von null annimmt, wenn genauso häufig optimale Eingriffe ausgeführt wurden, wie optimale Eingriffe bekannt waren (Anhang B). Ein negativer Wert des Maßes zeigt an, dass der Problemlöser mehr optimale Eingriffe kannte, als er ausgeführt hat, ein positiver Wert signalisiert eine bessere Regulation der Wissensanwendung, als aufgrund der Bekanntheit der Eingriffe zu erwarten war.

Das Gütemaß für die Regulation der Wissensanwendung wurde über alle Eingriffe hinweg berechnet, die ein Schüler bzw. eine Schülerin innerhalb der zur Verfügung stehenden Zeit bei der Bearbeitung der bis zu 22 Steuerungsaufgaben getätigt hat. Dies war notwendig, da zum einen nicht alle Schülerinnen und Schüler die gesamte Zeit benötigten, um die 22 Steuerungsaufgaben zu lösen, zum anderen aber auch manche Problemlöser es nicht schafften, alle 22 Aufgaben innerhalb der zur Verfügung stehenden

Zeit zu bearbeiten. Dadurch kann die Regulation der Wissensanwendung jedoch nur durch einen einzigen $\log_{(or)}$ -Wert bewertet werden. Bei der Erfassung der Regulation des Wissenserwerbs konnte der Wissenserwerbsprozess in mehrere Zeitintervalle unterteilt und somit mit mehreren $\log_{(or)}$ -Werten abgebildet werden. Das Regulationsmaß des Wissenserwerbs ist somit ein Verlaufsmaß, während das Regulationsmaß der Wissensanwendung diese nicht in Bezug auf ihren Verlauf, sondern hinsichtlich ihrer Gesamtgüte beurteilt.

4.3 Stichprobe und Durchführung

Die Erhebung mittels der computerbasierten Testverfahren fand an einem zusätzlichen dritten Testtag statt. In der Stichprobe des dritten Testtags waren 30 Schulen aus drei deutschen Bundesländern vertreten, in denen primär das dreigliedrige Schulsystem installiert ist¹. In den drei Landesstichproben waren die drei Schulformen Hauptschule, Realschule und Gymnasium entsprechend ihrem Anteil innerhalb des jeweiligen Landes proportional vertreten. Die Gesamtstichprobe des dritten Testtags setzte sich aus 22 Prozent Hauptschülern, 43 Prozent Realschülern und 35 Prozent Schülerinnen und Schülern des Gymnasiums zusammen. Insgesamt bearbeiteten 346 Schülerinnen und 342 Schüler das Raumfahrtspiel. Das durchschnittliche Alter lag bei 15,7 Jahren ($SD = 0,59$).

Auch wenn die Verteilung auf die drei Haupt-Schulformen ungefähr dem Anteil in den einbezogenen Bundesländern entspricht, kann für die Stichprobe keine Repräsentativität für die deutsche Schülerschaft angenommen werden. Aussagen über einzelne Subpopulationen oder gar Vergleiche verschiedener Länderpopulationen sind daher nicht sinnvoll. Die fehlende Repräsentativität ist jedoch kein Manko dieses Testteils von PISA 2000, da dieser ja nicht auf Vergleichsanalysen abzielte, sondern die Entwicklung und Evaluation eines neuen Testverfahrens im Auge hatte. Für solche Evaluationszwecke ist die Stichprobe sehr wohl geeignet.

Der dritte Testtag wurde an den Schulen zwischen einer und vier Wochen nach den ersten beiden Testtagen durchgeführt. Dafür bereiste ein mit 30 identisch konfigurieren Laptop-PCs ausgestattetes Team des Max-Planck-Instituts für Bildungsforschung die 30 Schulen. Die Erhebungen fanden als Gruppensitzung innerhalb eines Schulraums statt und wurden immer von zwei Testleiterinnen bzw. Testleitern geleitet. Innerhalb einer Sitzung bearbeiteten alle Schülerinnen und Schüler zunächst das Raumfahrtspiel und danach, nach einer kurzen Pause, das Ökologie-Planspiel.

Nachdem die Testleitung das Raumfahrtspiel gestartet hatte, wurden alle Instruktionen rein computerbasiert präsentiert. In einer kurzen Einleitung wurde die Bildschirmoberfläche mit den roten Schaltern und den blauen Anzeigen erläutert, ohne dass bereits Eingriffe getätigt werden konnten. Danach wurden die Schalter und Anzeigen der Rakete präsentiert, und die Schüler hatten fünf Minuten Zeit, dieses System zu explorieren. Alle dafür notwendigen Eingriffe konnten ausschließlich mit der Computermaus

vorgenommen werden. Nach dieser ersten Phase des Wissenserwerbs hatten die Schüler fünf Minuten Zeit, mit dem simulierten Planetenfahrzeug zu interagieren. Danach wurden ihnen sowohl die Rakete als auch das Planetenfahrzeug zur Exploration präsentiert. Direkt im Anschluss bearbeiteten die Schülerinnen und Schüler den Wissenserwerbtest, dessen zweimal acht Fragen auch direkt am Bildschirm gestellt wurden. Für die Beantwortung standen maximal 10,5 Minuten zur Verfügung. Anschließend wurden innerhalb von maximal 13 Minuten bis zu 22 Steuerungsaufgaben gestellt, bei denen die Schülerinnen und Schüler das zuvor erworbene und im Wissenserwerbtest unter Beweis gestellte Wissen so zielgerichtet anwenden mussten, dass das Raumfahrtspiel einen vorgegebenen Zielzustand annahm. Nach diesem Wissensanwendungstest hatten die Schülerinnen und Schüler eine längere Pause, bevor sie das Ökologie-Planspiel bearbeiteten.

Alle Aktionen, die ein Schüler oder eine Schülerin mit der Computermaus ausführte, wurden automatisch durch den Computer aufgezeichnet und in ein so genanntes *Logfile* geschrieben. Diese Verhaltensdaten sind die Grundlage für die folgenden Analysen.

4.4 Eigenschaften/Qualität des Messinstruments

4.4.1 Reliabilität

Für die Erfassung des Erfolgs des Wissenserwerbs wurden den Schülerinnen und Schülern nach der dreimal fünfminütigen Exploration acht Interpolations- und acht prospektive Aufgaben gestellt. Skalennalysen offenbarten, dass die jeweils erste Interpolations- bzw. prospektive Aufgabe eine geringe Trennschärfe aufweist, was vermutlich auf die Unbekanntheit des Itemformats zurückzuführen ist. Zu jeder dieser Aufgaben gab es im Test eine Parallelaufgabe. Für die Berechnung des Erfolgsmaßes für den Wissenserwerb wurden die jeweils erste Aufgabe und die dazu gehörige Parallelaufgabe ausgeschlossen. Die resultierende Skala weist mit Cronbachs $\alpha = .78$ bei zwölf Items eine zufriedenstellende interne Konsistenz auf und kann damit als reliabel eingestuft werden.

Die Skala, die sich aus allen 22 Steuerungsaufgaben für die Bewertung des Anwendungserfolgs bilden lässt, ist mit Cronbachs $\alpha = .92$ ebenfalls sehr reliabel.

Die beiden Maße für die Regulationsgüte des Wissenserwerbs bzw. der Wissensanwendung werden nicht klassisch als Summe oder Prozentwert berechnet. Insofern lässt sich die Reliabilität dieser Maße auch nicht klassisch im Sinne interner Konsistenz bestimmen. Analysen von Wirth (2004) zu Vorwissenseffekten und Intervallgrößen lassen jedoch den Schluss zu, dass auch die Reliabilitäten dieser Maße als ausreichend angesehen werden können.

4.4.2 Schwierigkeit

Für die Erfolgsmaße konnte bereits im Feldtest 1999 gezeigt werden, dass sie eine für die untersuchte Population angemessene Schwierigkeit aufweisen (Kap. 2). Aufgrund dessen, dass in der Haupterhebung 2000 der Erwerb von Wissen durch die drei fünfminütigen Explorationsphasen gegenüber dem Feldtest 1999 stärker vorstrukturiert und damit erleichtert worden war, verwundert es nicht, dass die Schülerinnen und Schüler im Durchschnitt einen höheren Erfolg beim Wissenserwerb aufweisen als im Feldtest (Tab. 4.1). Zum Teil ist dieser erhöhte Wert jedoch auch auf den Ausschluss des jeweils ersten Interpolations- bzw. prospektiven Items samt ihrer Parallelaufgaben zurückzuführen.

Die Skalen für den Erfolg beim interaktiven Erwerb und der zielgerichteten Anwendung von Wissen sind zwar etwas zu leicht. Sie weisen jedoch eine angemessene Streuung auf, weshalb nicht mit Deckeneffekten oder anderen Effekten eingeschränkter Varianz zu rechnen ist. Für die Regulationsmaße des Wissenserwerbs und der Wissensanwendung lassen sich Schwierigkeiten im klassischen Sinne nicht angeben.

	Mittelwert	Standardabweichung	Empirisches Minimum	Empirisches Maximum
Lernregulation	0,57	0,22	-0,10	1,27
Lernerfolg	0,64	0,22	0,00	1,00
Anwendungsregulation	-1,67	0,99	-5,02	0,58
Anwendungserfolg	0,66	0,26	0,00	1,00

Tabelle 4.1 Deskriptive Statistik der vier Maße dynamischen Problemlösens

4.3 Fairness

Wenn die vorgestellten Erfolgs- und Regulationsmaße gute Indikatoren für die Fähigkeit zum dynamischen Problemlösen sein sollen, dann ist zu fordern, dass ihr jeweiliger Wert von anderen Fähigkeiten oder Personenmerkmalen, die nicht Teil des dynamischen Problemlösekonstrukts sind, weitestgehend unabhängig ist. Zu solchen konstruktfernen Merkmalen zählen beispielsweise das Geschlecht oder auch die Erfahrung im und das Interesse am Umgang mit Computern (siehe auch Kap. 2). Für eine ausreichende Fairness ist zu fordern, dass Korrelationen zwischen Maßen des dynamischen Problemlösens und konstruktfernen Maßen möglichst niedrig, auf alle Fälle jedoch unter $r = .30$ liegen sollten. Die entsprechenden Koeffizienten statistisch bedeutsam gewordener Korrelationen sind in Tabelle 4.2 angeführt.

Es zeigt sich, dass Interesse an Computern und Computererfahrung in keinem praktisch bedeutsamen Zusammenhang mit den Leistungen beim computerbasiert erfassten dynamischen Problemlösen stehen. Ebenso scheint niemand aufgrund eingeschränkter Zugangs- oder Nutzungshäufigkeiten in seinen dynamischen Problemlöseleistungen

	Geschlecht ¹	Computererfahrung	Computerinteresse	Zugang		Nutzung	
				zu Hause	in der Schule	zu Hause	in der Schule
Lernregulation	0,13	0,12				0,09	
Lernerfolg		0,10			-0,11		
Anwendungsregulation	0,29	0,12	0,08	0,13		0,15	
Anwendungserfolg	0,26	0,19	0,14	0,15	-0,09	0,21	-0,09

¹ Kodierung: 0 = weiblich, 1 = männlich.

Es sind nur statistisch bedeutsame Korrelationen angegeben ($p < .05$).

Tabelle 4.2 Korrelationen der Maße dynamischen Problemlösens mit konstruktfernen Variablen

benachteiligt zu werden, auch wenn der Zugang und die Nutzungshäufigkeit eines Computers zu Hause in einem leichten positiven Zusammenhang mit der Anwendungsregulation und dem Anwendungserfolg stehen. Es könnten sich hier eventuelle Ähnlichkeiten zwischen dem Raumfahrtspiel und so manchem zu Hause gespieltem Computerspiel offenbaren. Für eine Prüfung dieser These stehen jedoch keine geeigneten Daten zur Verfügung. Zudem sind die Zusammenhänge praktisch von eher geringer Bedeutung.

Das Geschlecht steht in leichtem Zusammenhang mit der Wissensanwendung. Jungen erzielen in den Maßen für die Anwendungsregulation und den Anwendungserfolg tendenziell höhere Werte als Mädchen, was die Beobachtungen im Feldtest repliziert. Jedoch ist auch dieser Zusammenhang mit $r < .30$ recht gering, sodass nicht von einem bedeutsamen Geschlechtereffekt ausgegangen werden muss. Den Maßen kann demzufolge eine gewisse Fairness gegenüber Jungen und Mädchen zugesprochen werden.

4.5 Regulation des Wissenserwerbs

In diesem Kapitel beschäftigen wir uns etwas eingehender mit der Regulation des Wissenserwerbs und hierbei insbesondere mit den Unterschieden zwischen erfolgreichen und weniger erfolgreichen Problemlösern im Umgang mit dem komplexen und dynamischen System, wie es durch das Raumfahrtspiel simuliert wird. Diese Analysen dienen dem Zweck, Ansatzpunkte zur Förderung des selbstregulierten interaktiven Wissenserwerbs im Umgang mit komplexen und dynamischen Systemen aufzudecken.

In Abbildung 4.3 sind drei verschiedene durchschnittliche Regulationsverläufe dargestellt. Mittelt man über alle Schülerinnen und Schüler hinweg deren erreichte $\log_{(or)}$ -Werte der Regulation des Wissenserwerbs, ergibt sich für die fünf Minuten der zweiten Explorationsphase der Verlauf, der die Beschriftung „durchschnittlich erfolgreich“ trägt. Es zeigt sich, dass zu Beginn des Wissenserwerbsprozesses offensichtlich am meisten Wert auf das Identifizieren neuer Informationen gelegt wird. Dies ändert sich jedoch im

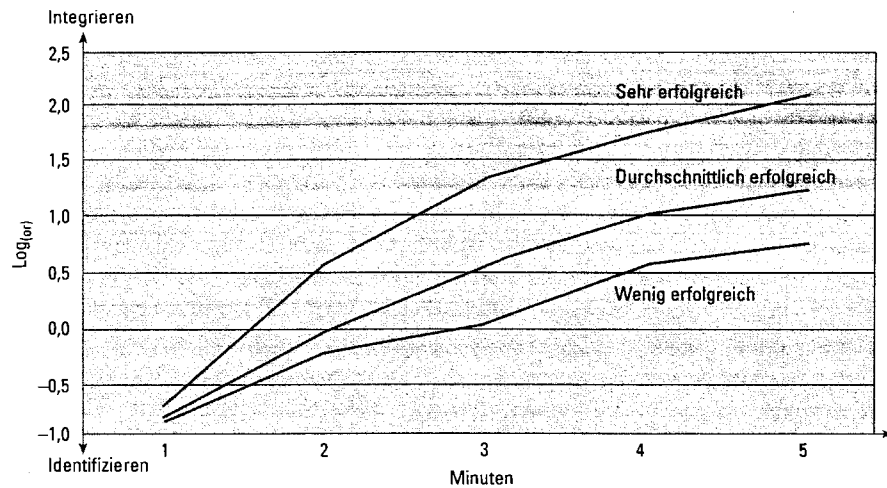


Abbildung 4.3 Unterschiedlich erfolgreiche Regulationsverläufe des Wissenserwerbs

Verlauf der Exploration, und das Integrieren identifizierter Informationen wird immer bedeutsamer für die Lernregulation.

Das Starten mit dem Identifizieren und der darauf folgende Wechsel zum Integrieren sind zwei Merkmale, mit denen sich generell der Regulationsverlauf des Wissenserwerbsprozesses beschreiben lässt (Wirth, 2004). Vergleicht man jedoch erfolgreiche mit weniger erfolgreichen Verläufen, so zeigen sich Unterschiede. In Abbildung 4.3 beschreibt der mit „sehr erfolgreich“ gekennzeichnete Verlauf die durchschnittliche Regulation der 10 Prozent besten Problemlöserinnen und Problemlöser, der „wenig erfolgreiche“ Verlauf gibt die durchschnittliche Regulation der 10 Prozent schwächsten wieder. Während sich bezüglich des Beginns keine Unterschiede zeigen – sowohl erfolgreiche als auch weniger erfolgreiche Problemlöser versuchen am Anfang verstärkt, neue Informationen zu identifizieren –, gehen erfolgreiche Problemlöser offenbar sehr bald und sehr stark dazu über, einmal identifizierte Informationen zu integrieren. Offenbar ist das Ausmaß dieses Wechsels vom Identifizieren und Integrieren ein differenzielles Merkmal erfolgreicher bzw. weniger erfolgreicher Problemlöserinnen und Problemlöser.

Um diesen Wechsel zu quantifizieren, wurde der Regulationsverlauf des Wissenserwerbs mithilfe dreier Faktoren im Rahmen eines latenten Wachstumskurvenmodells beschrieben (Abb. 4.4, siehe auch Anhang A). Der erste Faktor „Startwert“ gibt an, auf welchem Niveau der Prozess beginnt. Mit einem durchschnittlichen Faktorwert von $F_1 = -.79$ zeigt er an, dass im Durchschnitt mit dem Identifizieren von Informationen begonnen wird. Darin unterscheiden sich erfolgreiche Problemlöser nicht von weniger erfolgreichen. Der Faktor „Steigung“ quantifiziert den Wechsel vom Identifizieren zum

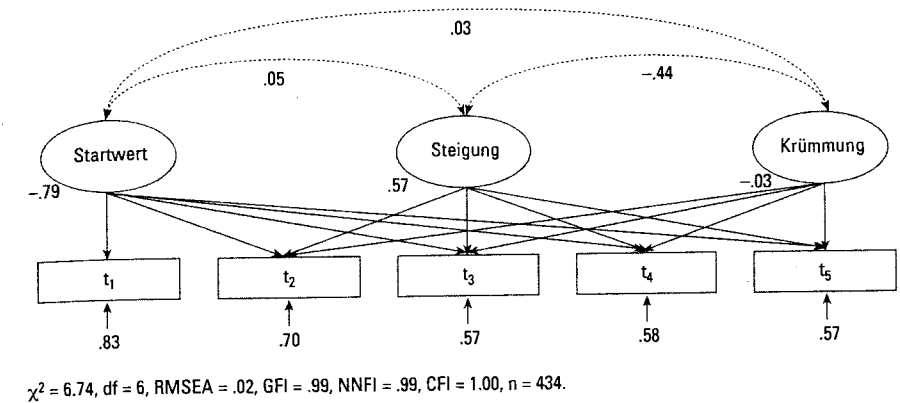


Abbildung 4.4 Latentes Wachstumskurvenmodell der Lernregulation

Integrieren. Im Durchschnitt nimmt dieser Faktor einen Wert von $F_2 = .57$ an. Erfolgreiche Problemlöser haben allerdings mit $F_{2(e)} = .80$ ($SD = .13$) einen statistisch bedeutsam höheren „Steigungs“faktorwert als weniger erfolgreiche Problemlöser ($F_{2(we)} = .43$, $SD = .17$). Der „Krümmungs“faktor gibt an, dass das Hinwenden zum Integrieren nicht durchgehend in gleicher Stärke zunimmt, sondern dass diese Zunahme mit der Zeit immer weiter nachlässt. Erfolgreiche und weniger erfolgreiche Problemlöserinnen und Problemlöser unterscheiden sich bezüglich dieses Faktorwerts nicht bedeutsam voneinander.

Erfolgreiche Verläufe der Regulation des Wissenserwerbs unterscheiden sich von weniger erfolgreichen Verläufen nicht hinsichtlich ihres Startniveaus oder hinsichtlich ihrer Krümmung. Sie unterscheiden sich jedoch sehr wohl hinsichtlich ihrer Steigung, also der wechselnden Ausrichtung der Regulation vom Identifizieren zum Integrieren. Aus testtheoretischer Perspektive ergibt sich daraus, dass der Wert des Steigungsfaktors als geeigneter Indikator der Regulation des Wissenserwerbs herangezogen werden kann. Dieser Indikator ist damit ein echtes Prozessmaß, das auf der Basis von Verhaltensdaten die Güte des Regulationsverlaufs zu erfassen vermag. Aus lehr-lernpsychologischer Perspektive zeigt sich in diesem Unterschied, dass es weniger erfolgreichen Problemlösern offensichtlich schwer fällt, einmal entdeckte oder generierte Informationen für spätere (Anwendungs-)Gelegenheiten festzuhalten. Der Aufwand, der für das Identifizieren betrieben wird, verpufft, weil einmal identifizierte Informationen wieder in Vergessenheit geraten. In späteren Situationen steht der Problemlöser damit wieder vor genau denselben Problemen, er bleibt sozusagen ein ewiger Problemlöser, während erfolgreiche Problemlöser im Umgang mit dem System dazulernen und die Steuerung des Systems immer mehr zur Routine wird. Instruktionale Maßnahmen zur Förderung dynamischer Aspekte von Problemlösekompetenz sollten demzufolge genau hier ansetzen und darauf

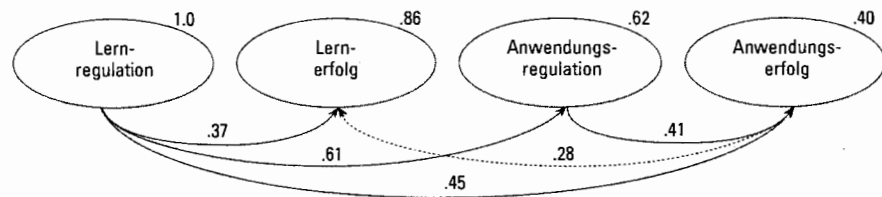
abzielen, dass einmal entdeckte oder generierte Informationen auch tatsächlich erworben werden.

4.6 Dynamisches Problemlösen

Die Regulation des Wissenserwerbs ist jedoch nur ein Aspekt des dynamischen Problemlösens. Dazu kommt auch die Fähigkeit, einmal erworbenes Wissen zielführend anzuwenden. Eine gute Wissenserwerbsregulation sollte zu einem erfolgreichen Wissenserwerb führen, genauso wie eine gute Anwendungsregulation einen hohen Anwendungserfolg nach sich ziehen sollte. Zudem ist anzunehmen, dass Personen, die sich bereits bei der Regulation des Wissenserwerbs nicht von Zufälligkeiten leiten lassen, sondern den Prozess selbstbestimmt regulieren, auch bei der Anwendung von Wissen einen hohen Grad an Selbstregulation aufweisen.

Wenn den hier vorgestellten vier Maßen für Aspekte dynamischen Problemlösens eine gewisse Konstruktvalidität bescheinigt werden soll, müssen sich mit ihnen diese Beziehungen auch empirisch nachweisen lassen. Ob dem so ist, wurde über Strukturgleichungsmodelle mithilfe von LISREL (Jöreskog & Sörbom, 1988, 1999) überprüft.

In Abbildung 4.5 ist das Modell mit der höchsten Anpassungsgüte dargestellt². Wie erwartet zeigt sich ein enger Zusammenhang zwischen der Güte der Regulation des Wissenserwerbs und der Höhe seines Erfolgs ($\beta = .37$). Einen ähnlich starken Effekt hat mit $\beta = .41$ die Anwendungsregulation auf den Anwendungserfolg. Für den Anwendungserfolg ebenso bestimmend ist aber auch die Fähigkeit zur Wissenserwerbsregulation. Dies kann vermutlich darauf zurückgeführt werden, dass erfolgreiche Problemlöser auch noch während der Bearbeitung der Steuerungsaufgaben des Wissensanwendungstests dazu in der Lage sind, dabei neu identifizierte Informationen in ihre persönliche Wissensstruktur zu integrieren. Die beiden Regulationsmaße stehen in einem engen Verhältnis zueinander ($\beta = .61$): Personen, die den Prozess des Wissenserwerbs zu regulieren wissen, sind offensichtlich auch dazu in der Lage, ihr bisher erworbenes Wissen optimal zielgerichtet anzuwenden. Gemeinsam klären diese beiden Maße 60 Prozent der Varianz des Anwendungserfolgs auf. (Direkt an den Ovalen der latenten



$\chi^2 = 50.11$, $df = 34$, $RMSEA = .04$, $GFI = .97$, $NNFI = .98$, $CFI = .99$, $n = 298$.

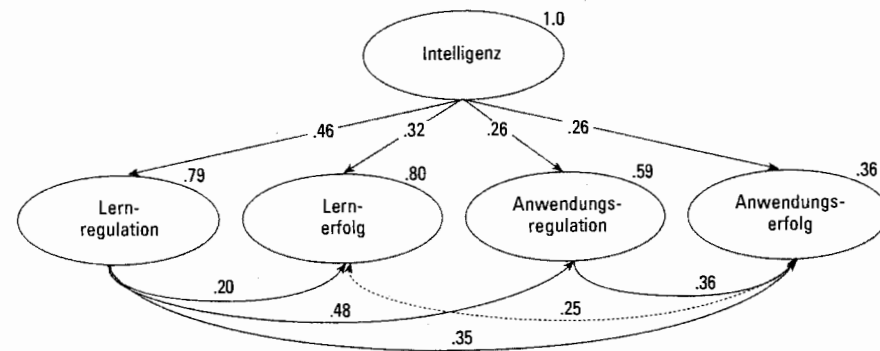
Abbildung 4.5 Struktur des dynamischen Problemlösens

Faktoren ist der Anteil an Varianz angegeben, der *nicht* durch das Modell aufgeklärt werden kann.)

Diese Analysen sind ein erster starker Hinweis darauf, dass die beiden Regulationsmaße genauso wie die beiden Erfolgsmaße als konstruktvalide Indikatoren für verschiedene Aspekte dynamischen Problemlösens anzusehen sind. Für weitere Validierungen ist jedoch zu prüfen, ob die damit erhobenen Aspekte dynamischen Problemlösens auch mit weiteren, konstrukt fremden Personenmerkmalen in zu erwartenden Beziehungen stehen. Dies wird im folgenden Abschnitt anhand des Personenmerkmals Intelligenz getestet.

4.7 Dynamisches Problemlösen und Intelligenz

Die Zusammenhänge zwischen Problemlösen und Intelligenz werden in der Literatur nicht einheitlich dargestellt und sind mit aller Wahrscheinlichkeit vermittelt durch das Ausmaß verfügbaren spezifischen Wissens („Elshout-Raaheim-Hypothese“; Elshout, 1987; Leutner, 2002; Raaheim, 1988; siehe auch Kap. 1 und 2). Es kann jedoch davon ausgegangen werden, dass Intelligenz einen durchaus positiven Einfluss auf die Fähigkeit zur Regulation des interaktiven Wissenserwerbs hat und dass auch die Bearbeitung der Aufgaben des Erfolgsmaßes für den Wissenserwerb eine gewisse Intelligenz voraussetzt (Kröner, 2001). Genauso ist im Sinne einer intelligenten Anwendung von Wissen ein Einfluss von Intelligenz auf die Anwendungsregulation und ihren Erfolg anzunehmen. Da jedoch davon ausgegangen werden kann, dass bei der Bearbeitung der Steuerungsaufgaben im Wissensanwendungstest mehr systemspezifisches Wissen zur Verfügung steht als während der vorangegangenen Exploration des Raumfahrtspiels, in der dieses Wissen ja zunächst erworben werden muss, dieses Wissen jedoch sicherlich noch



$\chi^2 = 45.08$, $df = 50$, $RMSEA = .00$, $GFI = .97$, $NNFI = 1.01$, $CFI = 1.00$, $n = 217$.

Abbildung 4.6 Struktur des dynamischen Problemlösens und Intelligenz

Gesamteffekte auf	Gesamteffekte von		Intelligenz
	Lernregulation	Anwendungsregulation	
Lernregulation			0,46
Lernerfolg	0,20		0,41
Anwendungsregulation	0,48		0,48
Anwendungserfolg	0,52	0,36	0,59

Tabelle 4.3 Gesamteffekte

nicht routinemäßig anwendbar ist, sollten im Sinne von Elshout und Raaheim die Effekte von Intelligenz auf die Wissensanwendung größer sein als auf den Wissenserwerb (vgl. entsprechende Überlegungen in Kap. 2).

Diese Annahmen wurden im Rahmen eines Strukturgleichungsmodells umgesetzt. Dafür wurde auf die Intelligenztest-Daten zurückgegriffen, die bei PISA 2000 am zweiten Testtag erhoben worden waren (KFT, verbale und figurale Analogien; Heller, Gaedicke, & Weinläder, 1985). Fit-Analysen ergaben, dass das in Abbildung 4.6 dargestellte Modell² am besten geeignet ist, die empirische Datenstruktur zu repräsentieren. Wie zu erwarten zeigt sich ein Einfluss der Intelligenz auf alle Aspekte dynamischen Problemlösens, wobei der direkte Einfluss auf die Wissenserwerbsregulation mit = .46 am höchsten ist.

Dies bedeutet jedoch nicht, dass die Intelligenz auf die Regulation des Wissenserwerbs einen höheren Einfluss hat als auf die Anwendung des Wissens, da Intelligenz nicht nur direkte, sondern auch indirekte, über andere Modellvariablen vermittelte Effekte hat. Die entsprechenden Gesamteffekte sind in Tabelle 4.3 angegeben. Hier zeigt sich, dass die Intelligenz mit den Variablen der Wissensanwendung in einem engeren Verhältnis steht als mit den entsprechenden Variablen des Wissenserwerbs. Einen nahezu ebenso hohen Einfluss auf die Anwendungsregulation und den Anwendungserfolg wie Intelligenz hat die Lernregulationsgüte. Dies ist insofern bemerkenswert, als die Lernregulationsgüte mithilfe des neu entwickelten Prozessmaßes erfasst wurde.

Damit lässt sich zusammenfassen, dass auch die Zusammenhangsanalysen zwischen Aspekten des dynamischen Problemlösens und der Intelligenz die erwarteten Effekte offenbaren, was als weiterer Beleg für die Konstruktvalidität aller vier Maße anzusehen ist.

4.8 Zusammenfassung

In diesem Kapitel wurde das dynamische Problemlösen vertiefend betrachtet. Dynamisches Problemlösen zeichnet sich durch zwei Anforderungen aus: Zum einen muss in der Interaktion mit dem unbekanntem, meist recht komplexen System Wissen über dessen Struktur und die Möglichkeiten seiner Steuerung erworben werden. Zum anderen

muss dieses Wissen für eine zielgerichtete Steuerung des Systems angewandt werden. Die Regulation des Wissenserwerbs erfordert ihrerseits wiederum das Verfolgen zweier Ziele: Zum einen müssen neue Informationen identifiziert werden, zum anderen müssen identifizierte Informationen so in die Wissensstruktur des Problemlösers integriert werden, dass er bzw. sie diese Informationen auch zu späteren Gelegenheiten abrufen und nutzen kann.

Für das Raumfahrtspiel wurden zusätzlich zu den beiden Erfolgsmaßen für den Wissenserwerb und der Wissensanwendung, die bereits im Feldtest 1999 erprobt worden waren (vgl. Kap. 2), innovative Regulationsmaße entwickelt und auf ihre Brauchbarkeit im Rahmen von *Large-Scale Assessments* getestet. Mithilfe des Regulationsmaßes für den Wissenserwerb konnte gezeigt werden, dass erfolgreiche Problemlöser sehr früh und sehr stark damit beginnen, einmal identifizierte Informationen zu integrieren, wohingegen weniger erfolgreiche Problemlöser das Integrieren nicht konsequent verfolgen und somit immer wieder vor denselben Problemen stehen. An dieser Stelle zeigt sich, wie eng das dynamische Problemlösen mit der Fähigkeit zum selbstregulierten Lernen assoziiert ist (Wirth, 2004). Eine gute Regulation des Wissenserwerbs zieht wiederum eine gute Anwendungsregulation im Sinne einer optimalen Nutzung erworbenen Wissens nach sich. Unter dem Einfluss von Intelligenz führt eine optimale Anwendungsregulation dann auch zu dem entsprechenden Anwendungs- bzw. Steuerungserfolg.

Die Regulationsmaße wurden mit dem Ziel entwickelt und eingesetzt, Ansatzpunkte zur Förderung dynamischen Problemlösens aufzuzeigen. Aus den hier vorgestellten Ergebnissen lässt sich ableiten, dass insbesondere der Wechsel vom Identifizieren zum Integrieren beim Wissenserwerb entscheidend für den Problemlöseerfolg ist. Insofern ist hier auch ein Hauptansatzpunkt zur Förderung zu sehen. Wenn es gelingt, Schülerinnen und Schülern die Wichtigkeit des Integrierens zu vermitteln, und wenn ihnen geholfen wird, ihre Wissenserwerbsregulation entsprechend zu optimieren, ist zu erwarten, dass damit die Fähigkeit dieser Schülerinnen und Schüler, mit unbekanntem komplexen und dynamischen Systemen kontrolliert umzugehen, deutlich verbessert wird.

Anmerkungen

¹ Aufgrund der Zugehörigkeit der 688 Schülerinnen und Schüler des dritten Testtags zu insgesamt nur 30 Schulen weist diese Stichprobe die hierarchische Struktur einer Klumpenstichprobe auf. Die anzunehmende Homogenität innerhalb einer Schule führt zu einer Unterschätzung des Standardfehlers. Das bedeutet, dass bei statistischen, hypothesenprüfenden Verfahren diese Unterschätzung ausgeglichen werden muss. Ein entsprechendes Verfahren schlägt Kish (1965) vor, mit dessen Hilfe ein „effektiver“ Stichprobenumfang berechnet wird, der dann dem statistischen Prüfverfahren zu Grunde gelegt wird (vgl. Wirth, 2004). Diese Korrekturmethode wurde bei allen in diesem Kapitel vorgestellten Analysen und Modellberechnungen genutzt. Da-

raus resultieren jedoch für unterschiedliche Analysen unterschiedliche Stichprobenumfänge.

- ² Der Einfachheit wegen ist in der Abbildung nur der „Steigungs“faktor als Indikator für die Regulationsgüte des Wissenserwerbs dargestellt und mit „Lernregulation“ beschriftet. Das berechnete Strukturgleichungsmodell, das der präsentierten Graphik zu Grunde liegt und für das die Anpassungsgütemaße angegeben sind, enthält jedoch alle drei latenten Wachstumsfaktoren zur Beschreibung des Regulationsverlaufs des Wissenserwerbs. Die Faktoren „Startwert“ und „Krümmung“ stehen in dem berechneten Modell aber in keiner bedeutsamen gerichteten oder ungerichteten Beziehung zu einem der anderen Faktoren.

5 Problemlösekompetenz – Ökonomisch und zugleich differenziert erfassbar?

In diesem Kapitel greifen wir erneut die Fragen nach der konvergenten und diskriminanten Validität der Maße des fächerübergreifenden Problemlösens auf, wie sie bereits im Kapitel 2 gestellt und dort zunächst auf der Basis der Erprobungsdaten beantwortet wurden. Damit wird unter anderem die Replizierbarkeit der dort dargestellten Befunde zur Kompetenzstruktur geprüft (siehe auch Klieme u.a., 2001; Wirth & Klieme, 2003). Darüber hinausgehend werden diese Fragen hier jedoch unter zwei zusätzlichen Perspektiven erörtert: Zum einen wird in Hinblick auf eine ökonomische Testanwendung geprüft, ob sich auch mit wenigen Testinstrumenten und einfach zu berechnenden Problemlösemaßen dieselbe Kompetenzstruktur abbilden lässt, wie sie bei der Erprobung, das heißt im so genannten Feldtest, unter Verwendung vieler Testverfahren aus unterschiedlichen Paradigmen der psychologischen Forschung zum Problemlösen aufgedeckt wurde. Dieser Aspekt der Testökonomie spielt gerade im Rahmen von *Large-Scale Assessments* und insbesondere beim Einsatz computerbasierter Testverfahren eine wichtige Rolle.

Zum anderen werden weiterführende Analysen der Kompetenzstruktur des Problemlösens vorgestellt. Während bei den Analysen des Feldtests zunächst zwischen analytischen und dynamischen Aspekten des Problemlösens unterschieden wurde, wird in den hier vorgestellten Analysen zusätzlich das dynamische Problemlösen differenzierter betrachtet. Vor allem werden die in den Kapiteln 1 und 2 eingeführten Fähigkeiten zum interaktiven Wissenserwerb und zur zielgerichteten Wissensanwendung auch empirisch voneinander unterschieden (vgl. auch Kap. 4). Dabei werden insbesondere die unterschiedlichen Zusammenhänge dieser beiden Teilkompetenzen mit den fachgebundenen Fähigkeiten im Lesen, in der Mathematik und in den Naturwissenschaften betrachtet.

Eckhard Klieme · Detlev Leutner
Joachim Wirth (Hrsg.)

Problemlösekompetenz von Schülerinnen und Schülern

Diagnostische Ansätze, theoretische
Grundlagen und empirische Befunde
der deutschen PISA-2000-Studie



VS VERLAG FÜR SOZIALWISSENSCHAFTEN

Bibliografische Information Der Deutschen Bibliothek
 Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie;
 detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

Inhalt

Einleitung	7
Detlev Leutner, Joachim Funke, Eckhard Klieme und Joachim Wirth	
1 Problemlösefähigkeit als fächerübergreifende Kompetenz	11
1.1 Einleitung: Fächerübergreifende Kompetenzen als Gegenstand von PISA	11
1.2 Der Kompetenzbegriff bei PISA	12
1.3 Problemlösen als Konzept und Forschungsgegenstand der kognitiven Psychologie	13
1.4 Problemlösen als Konzept und Forschungsgegenstand der differenziellen Psychologie	14
1.5 Analytische und dynamische Aspekte des Problemlösens	16
1.6 Selbstregulation beim dynamischen Problemlösen	18
Detlev Leutner, Joachim Wirth, Eckhard Klieme und Joachim Funke	
2 Ansätze zur Operationalisierung und deren Erprobung im Feldtest zu PISA 2000	21
2.1 Einleitung	21
2.2 Operationalisierung von Problemlösekompetenz	22
2.3 Empirische Erprobung im Feldtest	26
2.4 Zusammenfassende Diskussion	36
Eckhard Klieme, Johannes Hartig und Joachim Wirth	
3 Analytisches Problemlösen: Messansatz und Befunde zu Planung- und Entscheidungsaufgaben	37
3.1 Projektaufgaben: Konzeption und Auswertungsstrategie	39
3.2 Validität und Niveaustufen des analytischen Problemlösens	47
3.3 Analytische Problemlösekompetenz in Abhängigkeit von Schulform, Bundesland und Geschlecht	50
3.4 Fazit	52
Joachim Wirth und Joachim Funke	
4 Dynamisches Problemlösen: Entwicklung und Evaluation eines neuen Messverfahrens zum Steuern komplexer Systeme	55
4.1 Einleitung	55
4.2 Instrument und Leistungsmaße	56

1. Auflage September 2005

Alle Rechte vorbehalten
 © VS Verlag für Sozialwissenschaften/GWV Fachverlage GmbH, Wiesbaden 2005

Lektorat: Stefanie Laux

Der VS Verlag für Sozialwissenschaften ist ein Unternehmen von Springer Science+Business Media.
 www.vs-verlag.de



Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlags unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Umschlaggestaltung: KünkelLopka Medienentwicklung, Heidelberg
 Druck und buchbinderische Verarbeitung: MercedesDruck, Berlin
 Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier
 Printed in Germany

ISBN 3-531-14736-6