

웹 기반 심리 평가의 타당화*

곽 지 은 정 혜 원 곽 호 완[†]

경북대학교 심리학과

본 연구는 웹 기반 심리 평가의 타당도 검증을 위해 실험과 설문 평가에 대한 웹 기반과 실험실 기반 자료를 비교하였다. 먼저 연구 1에서는 웹 기반 실험의 타당도를 밝히기 위해 동일한 피험자들을 대상으로 주의력 검사 실험과제인 회귀억제(IOR) 과제와 연속 수행 과제(CPT)를 웹 기반과 실험실 기반으로 실시하였다. 상관표본 t 검증으로 분석한 결과, 이들 두 과제에 대한 웹 기반과 실험실 기반의 종속측정치들이 모두 통계적으로 유의한 차이가 관찰되지 않았다. 이는 웹 기반 실험의 수행 결과가 실험실 기반 실험의 결과와 동등함을 시사한다. 연구 2에서는 웹 기반 설문 평가의 타당도를 검증하기 위해 웹 기반과 지필형 설문에 중복 참여한 자료를 바탕으로 각 척도 평균을 비교하고, 상관을 알아본 결과, 지필형과 웹 기반 설문의 각 척도에 대한 평균의 차이가 유의하지 않았고, 두 방법을 통한 각 척도점수는 매우 높은 상관을 나타냈다. 이는 웹을 통한 설문 역시 지필형과 동등하게 해석할 수 있다는 것을 의미한다.

주요어 : 웹 기반 실험, 웹 기반 설문 평가, 타당도

* 이 논문은 2007년도 경북대학교 학술연구비에 의하여 연구되었음.

[†] 교신저자 : 곽호완, 경북대학교 사회과학대학 심리학과, 대구시 북구 산격동 1370번지
Fax: 053-950-5243, E-mail: kwak@knu.ac.kr

인터넷의 보편화로 인해, 전통적으로 실험실에서 행해지던 심리 평가가 최근에는 웹을 통해 구현되고 있는 추세이다. 치료자의 개입을 최소화하면서 내담자 스스로가 자기감찰(self-monitoring)을 통하여 문제를 조절하고 극복하도록 하는 자조치료(self-help treatment)를 바탕으로 한 온라인 상담과 심리치료가 점차 확산되면서(Richard, Klein, & Carlbring, 2003), 웹을 통해 질문에 응답하고, 결과를 전문가에게 전송하는 웹 기반 자기 보고형 심리 평가(web-based self-reporting test)의 개발 필요성 또한 증가하고 있다(Carlbring, Nilsson-Ihrfelt, Waara, Kollenstam, Buhrman, Kaldo, Soderberg, Ekselius & Andersson, 2005). 또 다양한 웹 기반 신경심리 실험 프로그램의 개발이 이루어지고 있는데, 특히 곽호완(2007)은 주의 과제에 대한 연구를 바탕으로 새로운 웹 기반 신경심리검사를 개발하였다. 그는 회귀억제 과제와 스트룹 과제 및 외생-내생 주의 과제의 세 가지 웹 기반 신경심리 실험 프로그램을 개발하여 성인 ADHD 성향군의 주의력 결함을 검토하였다. 그 결과, 회귀억제량, 스트룹 간섭량, 내생-외생 주의 단서 효과량, 각 과제의 평균 반응시간, 반응시간 편차 등의 실험 지표를 추출하였고, 이 중에서 회귀억제와 외생주의 단서 효과가 ADHD 증상 지표와 높은 상관성이 있음을 밝혔다. 또 판별분석을 이용하여 웹 기반 신경심리평가 지표들이 성인 ADHD 성향자들을 잘 판별해낼 수 있는지를 알아본 결과, 판별정확도가 95.7%로 나타났다.

실험실이 아닌 웹 기반으로 심리 평가를 실시할 경우, 시간과 비용을 절감할 수 있고, 다양하고 많은 피험자들의 자료를 비교적 용이

하게 수집할 수 있는 등 중요하고 실용적인 장점들이 많다(Reips, 2000, 2002). 그러나 이와 같은 장점에도 불구하고, 웹 기반 평가는 검사환경이나 시행방식이 검사결과에 영향을 미칠 수 있고, 참가자가 불성실하게 반응하게 될 가능성이 높아서 자료에 대한 신뢰성의 의심받기도 한다(Andersson, Kaldo-Sandstrom, Strom, & Stromgren, 2003; Buchanan, 2003). 이러한 이유로 웹 기반 검사의 경우, 실험실 기반 검사에 비해 검사 환경을 통제하기 어려워서 수집된 자료의 타당성이 문제가 될 수 있으므로(Brand, 2004; Brand & Hahn, 2003), 웹과 실험실에서 얻은 자료의 동등함을 밝힘으로써 웹 기반 연구의 타당성을 입증하고자 하는 연구가 이루어지고 있다(Musch & Reips, 2000; Pagani & Lombardi, 2000).

타당도에는 내적 타당도(internal validity)와 외적 타당도(external validity)의 두 가지 유형이 있다(Campbell, 1957). 내적 타당도는 독립변인과 종속변인의 인과관계를 추론할 때 사용하는 개념이고, 외적 타당도는 이러한 인과관계를 다양한 사람들과 다양한 상황에 일반화시킬 수 있는가를 뜻한다. 일반적으로 웹 기반 평가는 실험실 기반 평가에 비해 내적 타당도가 낮다. 이는 실험실 평가의 경우, 웹 기반 평가와 달리, 엄격하게 통제된 상황에서 혼입변인을 최소화할 수 있기 때문이다. 반면에 웹 기반 평가는 시간과 장비, 환경, 참가자가 다양하므로 일반화 가능성이 커서 외적 타당도가 높다(Birnbaum, 2000).

그러나 외적 타당도와 내적 타당도를 모두 설명할 때 전반적으로 실험실 평가가 웹 기반 평가에 비해 타당도가 높다(Brand, 2004). 심지

어 웹과 실험실 평가의 타당도가 전반적으로 거의 동등하다고 주장하는 연구자들조차 이 사실을 인정한다. 또 Cook, Campbell 및 Peracchio(1990)는 심리학 실험을 할 때 내적 타당도가 외적 타당도에 선행되어야 함을 강조하였다. 그래서 웹 기반 평가의 내적 타당도를 의심한 많은 연구자들이 이를 검증하기 위해 웹 기반과 실험실 기반 평가의 자료를 비교하였다(Birnbaum, 2001; Buchanan & Smith, 1999; Krantz, Ballard, & Scher, 1997; Reips, 2002).

서로 다른 두 유형의 자료를 비교하는 전형적인 방법은 상관분석과 회귀분석을 이용하는 것이다. 그러나 웹 기반과 실험실 기반 자료의 조건별 평균의 상관을 분석하는 이러한 방법은 다양한 조건의 많은 참가자 자료를 확보할 수 있을 때만 적합하다(Brand, 2004; Brand & Hahn, 2003). 더욱이 각 조건에서의 평균에 근거한 상관분석은 각 조건의 변량과 표본크기를 설명할 수 없어서 실험실과 웹 기반 자료간의 결정적인 차이를 애매하게 만들 수 있으므로, 보다 통계적 검증력이 크고 신뢰로운 비교 방법은 집단 간 평균 차이 검증을 실시하는 것이다(Brand, 2004; Brand & Hahn, 2003).

최근 국내에서도 웹 기반 심리 평가 시스템의 설계 및 구현에 관한 연구는 활발하게 이루어지고 있으나, 검사 결과 및 그에 대한 해석에 대한 연구는 많이 부족한 실정이다(서현주, 김선주, 주용석, 이수정, 이재호, 2001). 그리고 웹 기반 연구의 타당성을 입증하고자 하는 연구가 이루어졌으나(Musch & Reips, 2000; Pagani & Lombardi, 2000) 소수의 연구만이 웹 기반과 실험실 기반 실험 결과를 통계적으로

비교하고 있는바(예, Pagani & Lombardi, 2000; Krantz, Musch, & Naumer, 2000; Krantz, Ballard, & Scher, 1997), 웹 기반 심리평가의 타당도에 대한 더 많은 연구가 필요하다. 따라서 본 연구는 웹 기반 심리 평가의 타당도를 검증하기 위해, 실험과 설문 평가를 대상으로 웹 기반 자료의 타당도를 입증하고자 한다.

먼저 연구 1에서는 곽호완(2007)이 개발한 웹 기반 신경심리평가 중에서 주의력 검사인 회귀억제(inhibition of return, IOR) 과제와 연속 수행 과제(continuous performance test, CPT)를 대상으로 웹 기반 실험의 타당도를 검증하고자 한다. 회귀억제 과제는 억제적 주의의 결함을 파악하기 쉽게 설계되었고(곽호완, 1996), 비교적 평이하고 단순하여, 특정 피검자 유형에 국한되지 않고 억제적 주의 결함을 드러낼 수 있는 장점이 있다. 곽호완과 장문선(2007)에 따르면, 이 과제는 화면의 좌 또는 우측에 별표 자극이 제시될 때 가능한 한 빠르게 반응키(스페이스바)를 누르는 연속 반응 과제로, 연령이나 지능, 교육 수준, 숙련도 등에서 다양한 유형의 피검자들에게 적용 가능하다. 김영환, 진영선 및 곽호완(1995)도 알츠하이머형 노인성 치매의 주의 기능의 손상을 밝히기 위해 회귀억제 패러다임을 사용하였다.

연속 수행 과제는 ADHD 아동의 주의력 문제를 평가하기 위해 웹슬러 지능검사, 스트룹(stroop) 과제 등과 함께 주로 사용되어 왔다(곽호완, 장문선, 2007). 이 과제는 참가자로 하여금 화면을 관찰하게 하다가 빠른 속도로 점멸되는 글자나 숫자 자극으로 된 표적이 화면에 나타날 때 버튼을 누르도록 요구함으로써 지속적 주의 능력을 잘 드러내는 검사이다. 대

게 올바른 응답 횟수, 표적을 놓친 횟수, 방해 자극을 표적으로 잘못 인식한 횟수 등을 측정한다(정진영, 장문선, 곽호완, 2008).

연구 1에서는 회귀억제 과제와 연속 수행 과제를 각각 웹 기반과 실험실 기반으로 실시하고, 종속측정치들을 비교함으로써 두 가지 유형의 검사의 동등성을 검증할 것이다.

연구 2에서는 웹 기반 설문 평가의 신뢰도 검증을 위해, 성인 ADHD 평정척도인 CAARS-K, Barkley 유년기 ADHD 증상 척도의 ADHD 평정척도를 비롯하여, 정서적 측면과 관련하여 일반적으로 사용되는 Rosenberg 자존감 척도와 Beck 우울 척도, 그리고 대인관계에서의 부적응적 측면을 평정하는 한국형 대인관계 문제검사(Korean Inventory of Interpersonal Problem, 이하 KIIP)를 웹 기반과 지필형으로 실시하고, 각각에 대한 내적 합치도를 알아보고, 수집된 자료 중에서 두 방식 모두에 참여한 자료를 바탕으로 검사-재검사 신뢰도를 알아볼 것이다.

연구 1. 웹 기반 실험의 타당도 검증

연구 1에서는 곽호완(2007)이 개발한 웹 기반 실험의 타당도를 밝히기 위해 동일한 피험자들을 대상으로 주의력 검사인 회귀억제 과제와 연속 수행 과제를 실시하고, 그 결과를 상관표본 t 검증으로 분석할 것이다. 이 때 종속측정치들에 대한 집단 간 차이가 유의하지 않을 경우, 동등한 것으로 간주할 수 있다.

방 법

참가자 경북대학교 심리학과 대학원 신입생 14명을 실험에 참가하게 하였다.

장치 및 재료 웹 기반 실험은 곽호완(2007)이 개발한 웹 기반 신경심리검사 툴(<http://bh.knu.ac.kr/~kwak/psyttest>)을 이용하였다. 반면에 실험실 기반 실험을 위해서는 참가자들이 통제된 실험실에서 과제를 수행하여야 했다. 실험실에서 사용한 컴퓨터의 시스템 사양은 Microsoft Windows XP Home Edition Version 2002를 운영체제로 하는 IBM PC 호환 개인용 컴퓨터(Pentium(R) D CPU 3.40GHz, 504MB RAM)였고, 모니터는 17인치 LCD(HP L1706)였다.

실험 과제는 회귀억제 과제와 연속 수행 과제였다.

회귀억제 실험과제. 화면의 좌 또는 우측에 제시되는 '*' 자극을 탐지한 후 가능한 한 빠르게 반응키(스페이스키)를 누르는 연속 반응 과제였다. '+' 표시가 나타나고 500, 900, 1400msec의 특정 반응-자극 제시 시간차 후 '*' 자극이 제시되며, 참가자가 반응하면 자동으로 다음 시행으로 이어진다. 조작 변인은 이전 시행의 표적 위치와 현 시행의 표적 위치 동일성 조건(동일/상이 위치), 반응-자극 시간차(500, 900, 1400msec), 시행 블럭, 자극 제시 위치(좌/우) 등이다. 각 실험은 연습 시행을 포함하여 블럭 당 40시행씩 4블럭을 시행하였고 총 소요시간은 약 15분이었다.

연속 수행 과제. 먼저 화면의 중앙에 '+' 표시가 나타나면 눈을 이 곳에 고정시킨 다음,

화면의 좌 또는 우측에 잠깐 동안 제시되는 0에서 9까지 청색 숫자 들 중에서 '5' 가 아닌 자극(0, 1, 2, 3, 4, 6, 7, 8, 9)을 탐지한 후 가능한 한 빠르게 반응키(스페이스키)를 누르고, '5'이면 반응을 억제하는 과제로, 반응-무반응(go/no go) 절차를 포함시켜 오경보 반응과 누락 반응을 측정하였다. 자극 제시 시간은 200 msec, 자극 제시 시간 간격은 800msec이며, 각 실험은 연습 시행을 포함하여 블록 당 77시행씩 8블럭을 시행하였고 총 소요시간은 약 10분이었다. 조작 변인은 이전 시행의 표적 위치와 현 시행의 표적 위치 동일성 조건(동일/상이 위치) 등이다.

절차 및 자료분석 14명의 피검자로 하여금 웹 기반과 실험실 기반 회귀억제 과제와 연속 수행 과제를 모두 수행하게 하였다. 웹 기반 실험을 위해서는 시간과 장소의 통제 없이 스스로 과제를 수행하게 하였고, 실험실 기반 실험을 위해서는 정해진 시간에 동일한 실험실에서 실험자의 통제 하에 과제를 수행하게 하였다. 단, 연습효과 등을 고려하여 실시 순서를 역균형화하였다.

실험설계는 피험자내 설계로 동일한 피험자가 서로 다른 조건(웹 기반/실험실 기반)의 동일한 과제(회귀억제 과제와 연속 수행 과제)를 모두 수행하여야 했다. 회귀억제 과제의 종속 측정치는 회귀억제량, 평균 반응시간, 편차, 정반응률, 블록 효과, 좌우 자극 위치이었다. 연속 수행 과제의 종속 측정치는 회귀억제량, 평균 반응시간, 편차, 블록 효과, 누락 오류, 오경보 오류이었다.

회귀억제량은 동일 위치의 반응시간에서 상

이 위치의 반응시간을 뺀 값이다. 시행 블록 효과는 각 블록의 평균 반응시간의 기울기를 계산한 것으로, 피험자의 연습효과 또는 지속적인 주의결함을 보기 위한 것이다. 만일 연습 효과가 있다면 블록이 진행될 수록 반응시간이 빨라지는 부적 반응시간 기울기가 얻어질 것이며, 피로 또는 지속적인 주의 결함을 보인다면 정적 기울기가 얻어질 것이다. 웹 기반 실험의 타당도를 검증하기 위해 상관표본 *t* 검증을 실시하여 두 조건에서 얻어진 종속 측정치들을 비교하였다. 자료분석을 위한 통계 처리는 SPSS for Window 12.0 영문판을 사용하였다.

결과 및 논의

먼저 웹 기반과 실험실 기반 회귀억제 과제의 종속측정치들을 비교하기 위해 상관표본 *t* 검증을 실시하였다(표 1). *t* 검증 결과, 웹 기반과 실험실 기반 실험의 회귀억제량[$t(13) = 1.20, n.s.$], 평균 반응시간[$t(13) = -.57, n.s.$], 반응시간 표준편차[$t(13) = -.11, n.s.$], 유효 정반응률[$t(13) = 1.05, n.s.$], 블록 효과[$t(13) = .68, n.s.$], 좌우 자극 위치[$t(13) = .96, n.s.$]에서 모두 통계적으로 유의한 차이가 관찰되지 않았다. 이는 웹 기반 회귀억제 과제의 수행 결과가 실험실 기반 과제의 결과와 동등함을 의미한다.

다음으로 웹 기반과 실험실 기반 연속 수행 과제의 종속측정치들의 비교를 위해 상관표본 *t* 검증을 실시하였다(표 2). *t* 검증 결과, 웹 기반과 실험실 기반 실험의 회귀억제량[$t(13) = -1.19, n.s.$], 평균 반응시간[$t(13) = .22, n.s.$], 반응시간 표준편차[$t(13) = -.17, n.s.$], 블록 효과

표 1. 회귀억제 과제의 웹 기반과 실험실 기반 측정치들 간의 평균 차이 비교

실험지표	웹 기반 실험		실험실 기반 실험		t (df=13)
	평균	표준오차	평균	표준오차	
회귀억제량	17.50	4.32	11.79	4.49	1.20
평균 반응시간	454.79	20.51	462.57	13.43	-.57
반응시간 표준편차	70.98	5.80	71.56	4.64	-.11
유효 정반응률	99.36	.23	99.00	.23	1.05
시행 블록 효과	5.15	4.49	1.39	3.05	.68
좌우 자극 위치	.64	3.61	-4.21	3.73	.96

* $p < .05$.

표 2. 연속 수행 과제의 웹 기반과 실험실 기반 측정치들 간의 평균 차이 비교

실험지표	웹 기반 실험		실험실 기반 실험		t (df=13)
	평균	표준오차	평균	표준오차	
회귀억제량	5.21	1.59	7.14	1.97	-1.19
평균 반응시간	530.79	15.37	526.14	16.89	.22
반응시간 표준편차	71.16	3.82	72.63	7.18	-.17
시행 블록 효과	3.64	1.93	1.83	1.09	.86
누락 오류	.79	.24	.64	.20	.62
오경보 오류	28.07	3.92	20.79	2.90	1.26

* $p < .05$.

[$t(13)=.86$, *n.s.*], 누락 오류[$t(13)=.62$, *n.s.*], 오경보 오류[$t(13)=1.26$, *n.s.*]에서 모두 통계적으로 유의한 차이가 관찰되지 않았다. 이는 웹 기반 회귀억제 과제의 수행 결과가 실험실 기반 과제의 결과와 동등함을 의미한다.

를 검증하기 위해 웹 기반 설문과 지필형 설문 두 가지 모두에 중복 참여한 자료를 바탕으로 두 가지 자료수집 방식의 각 척도 평균을 비교하고, 상관을 알아보고자 하였다.

방 법

연구 2. 웹 기반 설문 평가의 신뢰도와 타당도 검증

연구 2에서는 웹 기반 설문 평가의 타당도

참가자 대구 소재 대학교에서 심리학 관련 과목을 수강 중인 대학생 813명을 대상으로 Conners ADHD 평정척도-한국판(CAARS-K),

Barkley 유년기 ADHD 증상척도, Beck 우울척도(BDI), Rosenberg 자존감 척도, 한국형 대인관계 문제척도가 포함된 설문지를 시행하였다. 설문지는 지필형(347명)과 웹 기반(466명)으로 실시하였고, 설문지 중 ADHD 반응 비일관성 지수(inconsistency index)가 8이상이거나 신뢰도가 의심되는 40부와 연령범위(18~29세)에 해당하지 않는 28부를 제외한 총 745부(남 342명, 여 403명)를 분석하였다. 연령 범위는 18-29세였으며, 평균 연령은 22.15세였다. 동일 피험자의 결과가 두 방식에 따라 다른지를 검토하기 위해 이들 중 지필형과 웹 기반 검사 모두에 중복 참여한 58명을 대상으로 분석을 실시하였다.

측정도구

Conners' 성인 ADHD 평정척도-한국판 (Conners' Adult ADHD Rating Scale- Korean: CAARS-K). Conners, Erhardt 및 Sparrow(1999a)가 개발한 성인 ADHD 평가척도를 김호영 등(2005)이 번안하고 타당화 한 것을 사용하였다. 총 66문항으로 이루어진 이 척도는 0(전혀 그렇지 않다)에서 3(매우(자주) 그렇다)점까지의 4점 리커트(likert) 척도로, Conners 등(1999b)이 요인분석을 통해 개발한 4개의 소척도와 DSM-IV 증상척도 2개, 및 ADHD 지수의 총 7개 소척도(부주의-기억 문제, 과잉행동, 충동성-정서적 불안정성, 자기개념의 문제, DSM-IV 부주의 증상, DSM-IV 과잉활동 및 충동성 증상)로 구성되어 있다. ADHD 지수는 Conners 등(1999b)의 연구에서 임상 집단과 정상 집단을 가장 잘 변별해주는 것으로 나타났던 문항들로 구성되었으며, 이 중 6문항은 다른 소척

도에도 중복되어 포함되는 문항이다. 그리고 일종의 타당도 지수로서 비일관성 지수를 제시하고 있는데, 이는 전체 문항 중 유사한 내용을 담고 있는 8쌍의 문항을 선정한 후 두 문항의 점수 차를 계산하여 합산한 값이며 8점 이상인 경우는 반응 일관성이 낮은 것으로 간주한다. 본 연구에서 내적 합치도(Cronbach's α)는 부주의 및 기억문제 .81, 과잉활동성 및 충동성 .82, 충동성 및 정서적 불안정성 .84, 자기개념 문제 .86, DSM_부주의 증상 .83, DSM_과잉활동 및 충동성 증상 .74, ADHD 지수 .77이었다.

Barkley 아동기 ADHD 증상척도. Barkley와 Murphy(1998)가 개발한 이 척도는 성인 ADHD 환자가 5-12세 시기 동안에 자신의 행동특성을 회상하여 증상을 0(전혀 그렇지 않았다)에서 3(매우(자주) 그렇다)까지의 점수 상에서 평정하도록 하는 자기보고식 척도이다. DSM-IV(1994)의 ADHD 증상으로 이루어진 총 18문항으로 구성되어 있으며, 홀수 문항은 부주의 증상을, 짝수 문항은 과잉행동-충동성 증상을 측정하며 각 문항의 점수를 합산하여 전체 ADHD 점수를 얻을 수 있다. 본 연구에서 내적 합치도(Cronbach's α)는 부주의 .82, 과잉활동 충동성 .79로 나타났다.

우울 척도. 우울을 측정하기 위해 Beck Depression Inventory(BDI)를 이영호와 송종용(1991)이 번안한 척도를 사용하였다. 이 척도는 Beck 등이 정서적, 인지적, 동기적, 생리적 영역을 포괄한 우울증상을 측정하기 위해 개발한 총 21개 문항의 자기보고형 검사이다.

총 21문항으로, 각 문항은 0점에서 3점까지 4점 척도로 평정한다. 점수 범위는 0점에서 63점까지이며 점수가 높을수록 우울함을 의미한다. 본 연구에서 사용된 BDI의 내적 합치도(Cronbach's α)는 .87이었다.

자아 존중감 척도. 자존감을 측정하기 위해서는 Rosenberg(1965)의 자존감 척도(Self-Esteem Scale; SES)를 이영호(1993)가 번안한 것을 사용하였다. 이 척도는 긍정적 자존감 5문항과 부정적 자존감 5문항 등 총 10문항으로 구성되어 있으며, 리커트 방식의 4점 평정 척도로서 점수가 높을수록 자존감이 높음을 의미한다. 본 연구에서의 내적 합치도(Cronbach's α)는 .86이었다.

대인관계 문제 척도. 대인관계 문제를 측정하기 위해서는 김영환, 진유경, 조용래, 권정혜, 홍상환, 박은영(2002)이 표준화한 한국형 대인관계 문제검사(Korean Inventory of Interpersonal Problem, 이하 KIIP)를 사용하였다. 이 척도는 개인이 일상생활에서 경험하는 여러 가지 대인관계 문제를 종합적으로 평가하고 가장 핵심적인 대인관계 역기능이 무엇인지 밝히고 성격문제를 선별하기 위한 자기보고형 검사이다. 이 척도는 8개의 원형척도(통제지배, 자기중심성, 냉담, 사회적 억제, 비주장성, 과순응성, 자기희생, 과잉관여)와 5개의 성격척도(대인적 과민성, 대인적 비수용성, 공격성, 사회적 인정욕구, 사회적 부족)로 구성되어 있으며, 각 척도는 전혀 그렇지 않다(0)에서 매우 그렇다(4)의 범위에서 평정할 수 있다. 본 연구에서 사용된 각 척도의 내적 합치

도(Cronbach's α)는 통제지배 .86, 자기중심성 .84, 냉담 .85, 사회적 억제 .90, 과잉관여 .76, 대인과민 .76, 대인비수용 .83, 공격성 .87, 인정욕구 .88, 사회적부족 .91이었다.

자료분석 웹을 통한 자료수집의 신뢰도를 알아보기 위해 지필형과 웹 기반 질문지 각각의 하위척도에 대한 내적합치도(Cronbach's α)를 SPSS 12.0을 사용하여 분석하였고, 또한 두 검사방법 모두에 중복 참여한 자료를 바탕으로 두 방식의 점수 간 피어슨 상관과 평균에 대한 차이검증을 실시하여 비교하였으며, 동일 피험자 내에서 결과의 동등성을 확인하고자 하였다.

결과 및 논의

ADHD 연구와 진단에서 자주 사용되는 질문지에 대한 응답의 타당도를 확인하기 위해 웹 기반 설문과 지필형 설문 두 가지 모두에 중복 참여한 자료를 바탕으로 자료수집 방식의 각 척도 평균을 비교하고, 상관을 알아보았다. 각 방식의 측정치에 대한 t 검증과 상관 분석 결과는 표 3에 제시된 바와 같다. 지필형과 웹 기반 설문의 각 척도에 대한 평균의 차이는 유의미하지 않았고, 두 방법을 통한 각 척도점수는 매우 높은 상관($r = .63 \sim .85$, $p < .01$)을 나타냈다. 이는 58명의 참가자들이 두 방식에서 비슷하게 반응을 했다는 것을 의미하고 웹을 통한 설문 역시 일반적인 방식인 지필형과 동일하게 해석할 수 있다는 것을 의미한다.

자기보고형 질문지에 대한 실험실 및 웹 기

표 3. 지필형 질문지와 웹기반 질문지의 평균비교와 상관(N=58)

	지필형	웹기반	<i>t</i>	<i>p</i>	<i>r</i>
	<i>M(SD)</i>	<i>M(SD)</i>			
Barkley 유년기					
유년기 부주의	6.53(3.77)	6.26(4.18)	.37	.71	.73**
유년기 과잉행동	5.97(3.72)	5.62(4.60)	.44	.66	.67**
CAARS-K					
부주의_기억	11.38(4.60)	11.47(5.30)	-.09	.93	.71**
과잉행동	11.97(5.16)	11.66(5.21)	.32	.75	.81**
충동_정서	11.43(5.21)	11.14(5.76)	.29	.77	.71**
자기개념	7.57(4.54)	6.83(4.57)	.88	.38	.84**
성인기 부주의	6.84(4.07)	6.91(3.89)	-.09	.93	.74**
성인기 과잉행동	6.45(3.44)	6.48(4.21)	-.05	.96	.72**
ADHD 지수	11.81(5.77)	10.45(5.73)	1.28	.20	.78**
Rosenberg 자존감					
BDI	8.69(6.15)	7.19(6.14)	1.31	.19	.71**
KIIP					
통제_지배	12.09(6.25)	11.52(5.87)	.51	.61	.64**
자기중심성	12.09(6.25)	11.52(5.87)	.51	.61	.64**
냉담	14.48(6.00)	14.22(6.52)	.22	.82	.75**
사회적 억제	14.29(7.60)	14.02(7.88)	.19	.85	.81**
자기희생	19.22(6.45)	18.38(7.13)	.67	.50	.85**
과잉관여	17.33(6.08)	17.12(6.16)	.18	.86	.63**
대인과민	16.03(5.43)	15.29(5.79)	.71	.48	.63**
대인비수용	16.03(6.00)	16.09(7.18)	-.04	.97	.70**
공격성	10.14(5.95)	9.45(5.22)	.66	.51	.65**
인정욕구	15.05(5.98)	14.91(6.11)	.12	.90	.74**
사회성부족	17.10(8.46)	16.83(8.87)	.17	.86	.71**

** *p* < .01.

반 검사의 신뢰성을 검증하기 위해서 두 가지 α 를 분석하였으며, 그 결과를 표 4에 제시하였다. 분석 결과 각 질문지의 하위 척도에 대

표 4. 지필형 질문지와 웹기반 질문지의 내적 타당도

	지필형(n=347)		웹기반(n=466)	
	M(SD)	Cronbach's α	M(SD)	Cronbach's α
Barkley ADHD 척도				
유년기 부주의	6.28(3.84)	.803	6.43(4.25)	.840
유년기 과잉행동	6.71(4.87)	.762	6.56(5.28)	.825
CAARS-K				
부주의_기억	10.61(4.72)	.793	10.71(5.26)	.830
과잉행동	10.31(5.07)	.804	10.53(5.70)	.841
충동_정서	10.54(5.30)	.833	10.17(5.73)	.853
자기개념	6.73(4.02)	.865	6.38(3.88)	.855
성인기 부주의	6.44(3.90)	.817	6.70(4.17)	.837
성인기 과잉행동	5.43(3.30)	.708	5.26(3.81)	.778
ADHD 지수	10.32(4.94)	.746	9.97(5.32)	.805
Rosenberg 자존감				
BDI	7.74(6.05)	.852	7.08(6.50)	.883
KIIP				
통제_지배	10.88(5.72)	.852	10.60(6.39)	.878
자기중심성	11.50(5.18)	.820	11.10(6.14)	.862
냉담	13.80(6.10)	.825	13.32(7.13)	.874
사회적 억제	13.15(7.27)	.890	13.64(8.20)	.900
비주장성	17.14(6.93)	.867	16.31(7.90)	.899
과순응성	17.83(6.69)	.846	17.20(7.54)	.874
자기희생	18.05(5.84)	.789	17.53(6.73)	.840
과잉관여	15.53(5.15)	.718	15.91(6.37)	.810
대인과민	14.58(4.92)	.719	14.33(5.95)	.803
대인비수용	14.56(5.68)	.805	14.19(6.80)	.858
공격성	9.18(5.43)	.861	8.82(6.06)	.887
인정욕구	13.67(5.35)	.859	13.60(5.94)	.884
사회성부족	15.98(8.37)	.899	16.00(9.39)	.919

** $p < .01$.

한 내적합치도는 지필형에서 .708-.899의 범위, 웹 기반에서는 .778-.919의 범위로 나타나, 두 방식 모두 .70 이상의 높은 내적합치도를 보였다. 두 방식의 내적합치도의 차이를 검증하기 위해 wilcoxon 순위검정을 실시하였고 그 결과, 전반적으로 웹 기반검사의 내적합치도가 의미있게 더 높은 것으로 나타났다($z(21) = -3.99, p < .001$). 이러한 결과는 웹 기반검사의 경우 미반응 항목이 있으면 서버로 제출되지 않고 미반응 항목위치로 가서 반응을 하게 하므로 반응하지 않은 항목이 많이 들어갈 수 있는 지필형 검사에 비해 더 나은 내적합치도를 보일 수 있음을 시사한다. 또한 웹 기반검사의 결과가 지필형과 유사하거나 혹은 더 우수한 검사임을 시사한다고 볼 수 있다. 다만, 각 방식의 평균(표준편차)에 관한 새로운 기준이 정립된다면 두 결과를 동일하게 해석할 수 있을 것이다.

종합논의

본 연구는 실험실 기반에 비해 검사 환경을 통제하기 어려워서 수집된 자료의 타당성이 문제가 될 수 있는 웹 기반 심리 평가의 타당도를 검증하고자 하였다. 이를 위해 심리 평가의 대표적인 두 유형인 실험과 설문 평가에 대한 웹 기반과 실험실 기반 자료를 비교하였다. 먼저 연구 1에서 웹 기반 실험의 타당도 검증을 위해, 기존의 웹 기반 신경심리평가(곽호완, 2007) 중에서 주의력 검사 실험 프로그램인 회귀억제 과제와 연속 수행 과제를 이용하여 동일한 대상에게 웹 기반과 실험실 기반으로 모두 실시한 다음, 상관표본 t 검증을 통

해 종속측정치들을 비교하였다. 그 결과, 이들 두 과제에 대한 실험실과 웹 기반 자료의 유의한 차이가 관찰되지 않았다. 따라서 웹 기반 실험의 결과가 실험실 기반 실험의 결과와 다르지 않으므로 실험이 가능하다고 볼 수 있다.

실험 2에서는 웹 기반 설문 결과를 기존의 지필형 설문 결과와 동일한 의미로 해석 가능한지 알아보기 위해 각 실시 방식에서의 내적합치도를 확인하고, 두 방식 모두에 참여한 동일 피험자 반응 평균의 차이와 점수 간 상관관을 알아보았다. 그 결과 Barkley 유년기 증상 척도와 CAARS-K 성인기 증상척도, Rosenberg 자존감 척도 및 Beck 우울 척도 모두에서 실시방식과 관계없이 .70 이상의 높은 내적합치도를 보였고, 동일 피험자 내에서의 반응 평균의 차이는 유의미하지 않았으며, 하위 척도 간 상관관이 .60이상으로 비교적 높은 상관관을 보였다. 이러한 결과는 웹을 통한 설문 반응 결과 해석의 호환가능성을 의미한다. 이에 반해 BDI-II에 대한 선행연구(Buchanan, 2003; Carbring et al., 2007)에서는 웹 기반 점수 평균이 유의미하게 높게 나타나 웹 기반 검사에 대한 새로운 기준을 성립할 필요가 있을 것으로 보았으나 Carbring 등(2007)에 따르면, 이는 검사 방식보다는 검사 시행 시점의 차이로 인한 것으로 두 시점 모두 동일한 방식으로 시행하게 되면 이 문제를 해결할 수 있을 것으로 보았다. 또한 반응 점수의 차이가 임상적으로 큰 영향을 미칠 정도는 아닌 것으로 나타났다. 본 연구에서도 의미 있는 차이를 보이지 않았다.

웹 기반 실험 연구는 기존의 현장 연구와

실험실 연구의 타당성 검증에도 유용할 뿐 아니라(Krantz & Dalal, 2000; Pohl, Bender, & Lachmann, 2002; Reips, 1997; Reips, Morger, & Meier, 2001), 웹이라는 매체를 통해서만 실현될 수 있는 새로운 연구 기회도 제공하였다. 첫째, 전통적인 실험실 연구에 비해 보다 많은 참가자들의 참여를 이끌어낼 수 있으므로, 수천 명의 참가자를 필요로 하는 실험도 가능하게 되었다. 둘째, 웹 기반 실험을 위해 개발된 프로그램은 전통적인 실험실에서도 충분히 사용할 수 있으나, 반대로 실험실용으로 개발된 프로그램은, 심지어 인터넷 기술이 사용된 것일지라도, 웹 기반 실험에 바로 적용하기는 어렵다(Reips, 1997). 마지막으로, 전통적인 실험법의 타당성을 저해하는 실험자 효과, 피험자 편향, 낮은 검증력 등의 문제점들도(Chapanis, 1970; Orne, 1962; Rosenthal & Fode, 1973; Rosenthal & Rosnow, 1969; Smart, 1966) 웹 기반 실험으로 어느 정도 해결 가능하다(Reips, 2000). 그러나 이러한 이점을 가진 웹을 성공적으로 이용하려면 방법적, 절차적, 기술적, 윤리적 사항 등을 고려하고 장점을 최대한 활용하려는 노력이 더 필요하다(Reips, 2000).

웹 기반 설문 평가 역시 다음과 같은 장점이 있다. 첫째, 실험실이나 지필형 검사에 비해 실시 요령 및 문항 제시 방법이 검사자에 관계없이 일정하게 통제될 수 있고, 각 문항에 대한 반응 시간과 반응 양상 등에 관한 정보를 수집할 수 있어서 문항이나 척도 타당화 연구에도 도움이 된다(서현주 등, 2001). 둘째, 기존의 검사 방법에 비해 시공간적 제약에서 자유롭다(Barak & Buchanan, 2003). 검사만을 위

한 별도의 장치나 도구가 필요 없고, 컴퓨터와 인터넷 등의 기본적인 환경이 갖추어진 곳이면 언제, 어디서든지 검사를 실시할 수 있으며, 결과 제출과 함께 코딩이 이루어지므로 빠르고 정확한 채점이 가능하므로 검사자와 수검자 모두 시·공간적 비용을 절감할 수 있다. 실제 임상 장면에서도 내담자에 대한 즉각적인 피드백 제공이 가능하며, 상담이나 다른 검사에 앞서 미리 실시할 수 있으므로 검사 시간을 단축할 수 있다. 뿐만 아니라, 적은 비용으로 많은 피험자에게 실시할 수 있어 대량의 자료 수집이 가능하고 학술적 연구에 있어서도 유용하게 쓰일 수 있다. 셋째, 수검자가 보다 친숙한 환경에서 검사에 임할 수 있으므로 장시간이 소요되고 때로는 긴장과 불안을 유발하는 검사(Tseng, Macleod, & Wright, 1997)를 보다 편안한 마음가짐으로 실시할 수 있도록 한다. 이는 수검자 자신의 현재 상태나 증상에 대해 집중할 수 있도록 하여 검사의 타당도 향상에 기여한다.

반면, 웹 기반 실험과 설문 평가와 같은 웹 기반 심리 평가의 경우, 개별적인 실시 환경이 적절히 통제되지 못할 경우 검사결과에 영향을 미치게 될 가능성이 높고, 컴퓨터에 익숙하지 않은 개인에게는 실시하기 어려운 단점이 있다. 즉, 모니터 화면의 크기나 화면 구성 등과 같이 검사 매체의 통제가 요구되는 실험은 경우에 따라 수검자의 반응이 달라질 수 있다. 그리고 일반적으로 컴퓨터를 통한 의사소통에서 개인들은 사회적 기준과 제약에 얽매이지 않는 경향이 있으며, 다소 탈억제(disinhibited)되는 경향이 있는 것으로 알려져 있다. 이는 수검자가 자기 자신에 대해 개방

하게 되어 응답의 진실성을 확보하는 장점이 되기도 하지만 동시에 사회적 불편감이나 주 호소하는 증상을 과소평가하게 되기도 한다 (Buchanan, Joinson and Ali, 2002).

본 연구에서는 비록 웹 기반 심리 평가의 타당도를 검증하여 실험실 기반 평가와 동등하다는 결과를 얻었으나, 이는 인터넷과 컴퓨터에 익숙한 대학생 집단을 대상으로 한 것이므로 전체 피검자 집단으로 일반화하기에는 제약이 있을 것으로 보인다. 따라서 추후 보다 다양한 연령대와 인구통계학적 특성을 보이는 집단에 대한 연구도 필요할 것으로 보인다.

참고문헌

- 곽호완, 장문선 (2007). 성인 ADHD 경향성에 대한 웹 기반 실험신경심리 연구: 회귀 억제, 스트룹 및 내생-외생 주의과제. 한국심리학회지: 임상, 26(4), 1039-1056.
- 곽호완 (1996). 회귀억제효과로 본 억제적 주의과정의 기제. 한국심리학회지: 실험 및 인지, 8(2), 133 -146.
- 곽호완 (2007). 신경심리평가 (Neuropsychological Testing). Retrieved October 10, 2007, from <http://bh.knu.ac.kr/~kwak/psytest/>
- 김영환, 진영선, 곽호완 (1995). 회귀억제효과로 본 알츠하이머 형 노인치매환자의 억제적 주의장애. 한국심리학회지: 실험 및 인지, 7(2), 131-142.
- 김영환, 진유경, 조용래, 권정혜, 홍상황, 박은영 (2002). 대인관계문제검사 매뉴얼. 학지사.
- 김호영, 이주영, 조상수, 이임순, 김지혜 (2005). 한국판 Conners 성인 ADHD 평정척도의 신뢰도 및 타당도 연구: 대학생 표본을 중심으로. 한국심리학회지: 임상, 24(1), 171-185.
- 서현주, 김선주, 주용석, 이수정, 이재호 (2001). 웹을 이용한 초등학생 심리검사 시스템의 설계 및 구현. 한국정보교육학회 논문지, 6(1), 228-235.
- 이영호 (1993) 귀인양식, 생활사건, 사건귀인 및 무망감과 우울의 관계: 공변량 구조모형을 통한 분석. 서울대학교 박사학위논문.
- 이영호, 송종용 (1991). BDI, SDS, MMPI-D 척도의 신뢰도 및 타당도에 대한 연구. 한국심리학회지: 임상, 10(1), 98-113.
- 정진영, 장문선, 곽호완 (2008). 성인 ADHD 성향군의 회귀억제와 반응억제 결함. 한국심리학회지: 일반, 27(1), 179-196.
- 채선희 (1995). 컴퓨터화된 개별적응검사와 지필식 검사에 있어서 문항 모수값의 동등성 검증, 교육평가연구, 8(2), 145-160.
- Andersson, G., Kaldö-Sandström, & Strömgren, T. (2003). Internet administration of the Hospital Anxiety & Depression Scale in a sample of tinnitus patients. *Journal of Psychosomatic Research*, 55, 259-262.
- Barak, A. & Buchanan, T. (2003). Internet-Based Psychological Testing and Assessment. In Kraus, R., Zack, Jason., Stricker, G.(Ed.), *Online Counseling: A Handbook for Mental Health Professionals*. San Diego: Academic Press.

- Barkley, R. A., & Murphy, K. R. (1998). *Attention-Deficit Hyperactivity Disorder: A Clinical workbook*. New York: Guilford Press.
- Birnbaum, M. H. (2000). Validity of web-based psychological research. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 35-60). San Diego: Academic Press.
- Birnbaum, M. H. (2001). Introduction to behavioral research on the Internet. In M. H. Birnbaum (Ed.), *Introduction to behavioral research on the Internet* (pp. 1-12). Upper Saddle River, NJ: Prentice-Hall.
- Brand, A. (2004). A web experiment based enquiry into the verbal overshadowing effect. Retrieved April 4, 2008, from [http://www.ipsychempts.com/Brand_\(2004\).pdf](http://www.ipsychempts.com/Brand_(2004).pdf)
- Brand, A. (2004). A web experiment based enquiry into the verbal overshadowing effect. Retrieved April 4, 2008, from [http://www.ipsychempts.com/Brand_\(2004\).pdf](http://www.ipsychempts.com/Brand_(2004).pdf)
- Brand, A., & Hahn, U. (2003). Evaluating the validity of internet experiments: verbal overshadowing as a case study. Retrieved April 4, 2008, from [http://www.ipsychempts.com/Brand_and_Hahn_\(2003\).pdf](http://www.ipsychempts.com/Brand_and_Hahn_(2003).pdf)
- Buchanan, T. (2003). Internet-based questionnaire assessment: appropriate use in clinical contexts. *Cognitive Behavior Therapy*, 32 (3), 100-109.
- Buchanan, T., & Smith, J. L. (1999). Research on the internet: validation of a World-Wide-Web mediated personality scale. *Behavior Research Methods, Instruments and Computers*, 31, 565-571.
- Buchanan, T., Joinson, A. N., & Ali, T. (2002). Development of a behavioural measure of self-disclosure for use in online research. *German Online Research' 02*. Hohenheim, Germany.
- Carlbring, P., Brunt, S., Bohman, S., Austin, D., Richards, J., Öst, L., & Andersson, G. (2007). Internet vs. paper and pencil administration of questionnaires commonly used in panic/agoraphobia research. *Computers in Human Behavior*, 23, 1421-1434.
- Carlbring, P., Nilsson-Ihrfelt, E., Waara, J., Kollenstam, C., Buhrman, M., Ekselius, L., & Andersson, G. (2005). Treatment of panic disorder: live therapy vs. self-help via the internet. *Behaviour Research and Therapy*, 43 (10), 1321-1333.
- Chapanis, A. (1970). The relevance of laboratory studies to practical situation. In D. P. Shultz (Ed.), *The science of psychology: critical reflections*. New York: Appleton Century Crofts.
- Conners, C. K., Erhardt, D., Epstein, J. N., Parker, J. D. A., Sitarenios, G., & Sparrow, E. (1999b). Self-ratings of ADHD symptoms in Adults I: Factor structure and normative data. *Journal of Attention Disorder*, 3, 141-151.
- Conners, C. K., Erhart, D., & Sparrow, E. P. (1999a). *Conners' adult ADHD rating scales, technical manual*. New York: Multi-Health Systems.

- Cook, T. D., Campbell, D. T., & Peracchio, L. (1990). Quasi-experimentation. In M. D. Dunnett & L. M. Hough(Eds.), *Handbook of industrial and organizational psychology* (2nd ed., 1, pp. 491-576). Palo Alto, CA: Consulting Psychologists Press.
- Krantz, J. H., & Dalal, R. S. (2000). Validity of web-based psychological research. In M. H. Birnbaum (Ed.), *Psychological experiments on the internet* (pp. 35-60). San Diego, C.A.: Academic Press.
- Krantz, J. H., & Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review*, 107, 852-884.
- Krantz, J. H., Ballard, J., & Scher, J. (1997). Comparing the results of laboratory and World-Wide Web samples on the determinants of female attractiveness. *Behavior Research Methods, Instruments and Computers*, 29, 264-269.
- Musch, J., & Reips, U. D. (2000). A brief history of Web experimenting. In M. H. Birnbaum (Ed.), *Psychological experiments on the internet* (pp. 61-87). Orlando, FL: Academic Press.
- Orne, M. T. (1962). On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. *American Psychologist*, 27, 959-962.
- Pagani, D., & Lombardi, L. (2000). An intercultural examination of facial features communicating surprise. In M. Birnbaum (Ed.), *Psychological experiments on the internet* (pp. 169-194). Orlando, FL: Academic Press.
- Pohl, R. F., Bender, M., & Lachmann, G. (2002). Hindsight bias around the world. *Experimental Psychology*, 49 (4).
- Reips, U. D. (1997). Das psychologische experimentieren im Internet [Psychological experimenting on the Internet]. In B. Batinic (Ed.), *Internet for psychologen* (pp. 245-265). Gotingen, Germany: Hogrefe.
- Reips, U. D. (2000). The web experiment method: advantages, disadvantages, and solutions. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 89-114). Orlando, FL: Academic Press.
- Reips, U. D. (2002). Standards for internet-based experimenting. *Experimental Psychology*, 49(4), 243-256.
- Reips, U. D., Morger, V., & Meier B. (2001). Funfe gerade sein lassen: listenkontexteffekte beim Kategorisieren [Letting five be equal: list context effects in categorization]. Retrieved April 7, 2002 from <http://www.psychologie.unizh.ch/genpsy/reips/papers/remome2001.pdf>
- Richard, J., Klein, B., & Carlbring, P. (2003). Internet-based treatment for panic disorder. *Cognitive Behavior Therapy*, 32, 125-135.
- Buchanan, T. (2002). Online assessment: desirable or dangerous? *Professional Psychology: Research and practice*, 33, 148-154.
- Rosenberg, M. (1965). *Society and adolescent self-image*. Princeton, HJ. Princeton University press.

- Rosenthal, R., & Fode, K. L. (1973). The effect of experimenter bias on the performance of the albino rat. *Behavioral Science*, 8, 183-189.
- Rosenthal, R., & Rosnow, R. L. (1969). The volunteer subject. In R. Rosenthal & R. L. Rosnow (Eds.), *Artifact in behavioral research*. New York: Academic Press.
- Smart, R. (1966). Subject selection bias in psychological research. *Canadian Psychologist*, 7a, 115-121.
- Tseng, H., Macleod, H. A., Wright, P. (1997). Computer anxiety and measurement of mood change. *Computers in Human Behavior*, 13, 305-316.
- 1 차원고접수 : 2008. 8. 2.
최종게재결정 : 2008. 12. 18.

Validation of Web-based Psychological Assessments

Ji-Eun Kwahk Hye-Won Jeong Ho-Wan Kwak

Department of Psychology, Kyungpook National University

This study compared data from web-based psychological experiments and self-reporting questionnaires with data from laboratory experiments and paper-pencil tests to verify the validity of web-based psychological assessment. In study 1, a total of 14 subjects participated in a comparative experiment, in which each participant performed both laboratory and two types of web-based attention tasks(IOR and CPT). The results of paired-sample *t*-test, there were no significant differences between the measurements for two different formats. This result may show the equivalence of the results from the two types of experiments(lab-based and web-based). In study 2, we examined paired-sample *t*-test and test-retest pearson correlation of each self-reported questionnaire subscale to evaluate if there were statistically significant differences between paper-pencil and web-based formats. The results revealed that mean subscales for two formats were not significantly different and subscale scores were highly correlated with each other. This results also suggested that web-based questionnaire data can be dealt equivalently with paper-pencil administered data.

Key words : web-based experiment, web-based questionnaire, validity