

노인용 이야기 회상 검사의 표준화 연구

안 효 정

강남성모병원 정신과

최 진 영[†]

서울대학교 심리학과

본 연구는 노인용 이야기 회상 검사를 제작하여 신뢰도와 타당도를 확인하고 검사 수행에 영향을 미치는 인구 통계학적 변인을 검토하여 이에 따른 규준을 마련하는데 목적이 있다. 규준 집단은 서울, 경기, 예산, 마산, 대구, 당진 등의 지역에 거주하는 만 55~89세 사이의 노인 273명으로 구성되었다. 검사 수행에 미치는 연령, 교육, 성별 등의 인구통계학적 변인의 효과를 분석한 결과, 교육이 매우 유의미한 요인으로 확인되었다. 규준은 즉각적 자유회상은 0~9년과 10년 이상인 두 교육 집단과 그에 따른 성별집단으로 세분화하여 4집단의 평균과 표준편차의 규준을 제시하였고, 지연 자유회상과 재인 회상은 55~74세와 만 75세~89세의 두 연령 집단으로 나눈 후 다시 0~9년과 10년 이상인 두 학력 집단으로 나누고 또 각기 성별 집단으로 다시 세분화한 8집단의 평균과 표준편차를 제시하였다. 그 외 이야기 회상 검사의 여러 수행 측정치들을 가지고 시간의 경과에 따른 기억 수행의 저하, 단서에 의한 재인 회상 수행 향상, 양호한 보유 능력 등 정상 노인들의 다양한 학습 및 기억수행 특징들을 관찰하였다. 노인 규준 자료는 임상 장면에서 치매와 신경학적 손상 환자들의 학습과 언어적 기억 및 전반적인 인지 능력에 대한 효율적인 판별에 기여하리라 기대된다.

주요어 : 이야기 회상 검사, 기억, 치매, 신경심리학, 알츠하이머 질환

* 본 연구는 첫 번째 저자의 석사학위 논문을 바탕으로 그 이후 추가 자료 수집을 통합하여 집필한 것임
이 연구는 과학기술부의 뇌신경정보사업의 지원으로 수행되었음.

저자들은 자료 수집에 기여한 김명진, 김호영, 송호정에게 감사의 뜻을 전함.

[†] 교신저자(Corresponding Author) : 최 진 영 / 서울대학교 심리학과 / 서울특별시 관악구 신림동 산 56-1
FAX : 02-880-6428 / E-mail : jychey@snu.ac.kr

노인기 행동 질환 중 최고의 발병률을 차지하는 치매는 노인 인구의 증가와(통계청, 1998) 더불어 지속적으로 늘어날 것이 예상된다(보건복지부, 1998). 치매는 정상적인 지적 발달을 이룬 성인이 어느 시점에 두 세 개 이상의 지적 기능이 감퇴되고 그 증상들이 지속되는 증후군이며(Cummings & Benson, 1992) 여러 종류의 질환이 원인이 될 수 있다. 이 중 알츠하이머 질환으로 인한 치매(Dementia of the Alzheimer Type: DAT)는 여러 가지 인지 기능이 심각하게 손상되는 퇴행성 치매로 전체 치매의 절반 정도를 차지하여 유병률이 가장 높은 치매이다. DAT의 임상적 특징인 인지적 손상의 측정을 위한 신경심리 검사들이 많이 개발되었는데 이 중에서도 기억 기능을 측정하는 검사들의 유용성이 확인되고 있다. 기억 장애는 DAT 초기부터 관찰되는 핵심 인지 결함으로 기억 측정치들이 다른 신경 심리 측정치들이나 정신 상태 검사 수행치들 보다 DAT를 변별하는데 효용성이 높은 것으로 보고되고 있다(Christensen, Hardzi-Pavlovic, & Jacomb, 1997).

기억 및 학습 능력 등 인지 기능들은 연령이 증가함에 따라 감퇴되는 경향이 있는데, 특히 노년기에 이르러 두드러지게 관찰된다(Lezak, 1995; Zec, 1993). 기억의 감퇴가 경미한 65세 이상의 치매 환자를 평가하는 경우, 정상 노화에 따른 기억 감퇴와 치매로 인한 병리적 기억 장애를 변별해주어야 한다(Kaszniak, 1986; Knopman & Ryberg, 1989; Loring, Lee & Meador, 1989). 또한 노인기에 빈발하는 뇌혈관 질환, 당뇨, 갑상선 질환 및 비타민 결핍 등 다른 질병으로 인한 기억 저하와도 변별해 주어야 한다. 치매 유병률이 우리 보다 먼저 증가한 선진국의 치매 연구에서는 변별이 가장 어려운 DAT 초기 단계와 정상 노인을 구분하는 데 언어 기억 과제의 유용성이 일관되게 관찰되고 있고 이중 문장들을 기억하

도록 하는 이야기 회상 검사의 유용성도 계속 확인되고 있다(Eslinger, Benton & Damasia, 1984; Jones, Benton, Paulsen & Tranel, 1992).

이야기 회상 검사는 단순한 단어 목록 보다 좀 더 복잡한 언어 자극인 문장들을 기억하게 하므로 주의 폭, 학습능력, 언어적 이해, 및 기억의 세부적인 과정인 부호화, 저장, 인출 과정 등에 대한 평가가 가능하고 단어 및 문장의 의미가 기억력에 미치는 영향을 동시에 측정할 수 있다(Lezak, 1995). 이 검사는 주인공에게 일어나는 사건을 논리적으로 기술한 문장들로 구성되어 있다. 이야기는 보통 정상인의 주의 폭 보다 긴 문장들, 주로 한 문단을 들려주고 이후에 회상하게 하는데 일반적으로 즉각 회상과 지연 회상의 절차가 있다. 이야기 회상 검사 중 대표적인 Wechsler Memory Scale의 최신판(Wechsler Memory Scale-III)의 Logical Memory(이하, LM) 소검사는 재인 과제도 추가되어 있다(Lezak, 1995; Wechsler, 1997). 이 검사의 수행은 사고를 얼마나 잘 조직화(organization)할 수 있는가를 반영하여(Lezak, 1995, 참조), 이야기를 시간의 흐름과 논리적 전개에 따라 전개하는 능력 및 관리 기능(executive function)을 필요로 하여 뇌의 전반적인 인지 기능을 필요로 하는 검사이다(O'Donnell, Drachman, Lew & Swearer, 1988).

DAT 환자들과 정상 노인들의 기억 수행을 비교할 때 유용한 이야기 회상 검사의 기억 측정치들로는 기억 과정 중 부호화 결함을 측정할 수 있는 즉각적 자유회상의 평균점수와 정보의 파지 및 저장을 측정하는 지연 자유회상의 평균 점수, 그리고 보유율(saving score or percentage retained scores)이 있다. 이러한 이야기 회상 검사의 측정치들은 기억 능력에 대한 양적인 정보뿐만 아니라 기억의 인지적 과정에 대한 질적 정보를 제공해 준다(Dikmen, McLean, Mechamer &

Temkin, 1990; Webster, Godlewski, Hanley & Sowa, 1992). 특히, 새로운 자극을 학습하여 지연된 시간 후에 회상하는 정도를 측정하는 이차적 언어 기억(secondary verbal memory) 검사로 초기 DAT 감별에 민감한 도구인 것으로 밝혀졌다(Storandt, Berg, Botwinick, Danziger & Hughes, 1984). 초기 치매 단계의 환자들을 대상으로 여러 신경 심리 검사의 수행을 평가한 연구에서 다른 신경 심리 검사와는 달리 LM 소검사에서 환자들이 일관되게 수행 저하가 관찰되었고 수행 차이도 가장 심하였다. 이는 DAT 환자에게 있어서 가장 먼저 출현하는 기억 장애를 이야기 회상 검사가 예민하게 측정하여 정상인과 초기 치매 환자를 잘 구별해 줄 수 있다는 것을 시사해 주는 결과라 하겠다(Morris, Ball, Berg, Grant, Mekeel, Price, Rubin, & Storandt, 1991).

나이에 따른 인지 결함을 보이는(Age-related cognitive deficit: ARCD) 노인들의 종단 연구 결과, 이 노인들 중 이후에 치매로 발전되는 사람들과 그렇지 않은 사람들로 판별하는데 WMS의 LM이 가장 유용한 검사로 확인되었다(Christensen et al., 1997). 다른 기억 검사들과 마찬가지로 치매 내의 변별 진단과 치매의 진행 정도를 판단하는 데에도 우수한 도구로 알려져 있다(Storandt et al., 1984). LM을 비롯한 이야기 회상 검사에서 사용되는 주요 지표 중 하나인 기억의 보유율은 즉각적 회상 수행 대비 지연 회상의 수행인데, 이 지수는 정상 노인 집단, 초기의 DAT 집단, 및 헌팅턴 집단들을 감별하는데 유효한 것으로 밝혀졌다(Welsh, Butters, Heyman, Hughes & Mohs, 1992). 또한, 헌팅턴 환자(Huntington's disease: HD)들과 초기의 DAT 환자들을 비교하는 연구에서는 LM의 보유율이 가장 변별력이 높은 측정치로 보고되었다(Kramer, Brandt, Delis & Levin, 1988). 또 다른 연구에서는 보유율에 의한 두 집단 간

의 변별이 질병의 아주 초기 단계까지도 가능하다고 보고하고 있다(Delis, Butters, Cermak, Kramer, Massman & Salmon, 1991).

이야기 회상 검사는 일상 생활에서 기억 기능을 잘 반영하기에(Lezak, 1995) 저학력자가 많은 한국 노인들의 치매 평가에서 중요한 역할을 할 수 있다. 현재 국내 임상 장면에서 사용되는 이야기 회상 검사로는 번역한 WMS-R(1987)의 LM이 있는데 아직 한국판으로 재개발하지 못하였기 때문에 몇 가지 점에서 문제점을 안고 있다. 신경심리 검사에서의 수행치의 의미는 규준에 의해 결정되는데 아직 국내 인구를 대상으로 수집된 규준이 없어 임상가들이 검사 점수를 해석하기 어렵다. 또한 번역한 언어 자극의 경우 원 검사가 개발될 당시 고려되었던 언어 자극의 특성을 잃기 때문에 기억 회상에 영향을 미치는 중요한 요인 중 하나인 자극의 친숙성이 떨어져 결과적으로 기억 검사로서의 타당성을 저하시킬 가능성이 높다(김정기, 강연옥, 1998; 최진영, 이소애, 1997). 문장으로 엮어진 이야기의 자극으로서의 친숙성을 결정하는 것은 어휘 빈도와 이야기 주제의 친숙성인데(Lezak, 1995), 모든 어휘의 사용 빈도는 언어와 문화에 따라 다르며 이야기 주제의 친숙성도 문화권마다 차이가 있다. 영어권에서 개발된 검사를 번역하거나 한국형을 만들 때 이러한 점을 고려하지 않으면 원 검사와 전혀 다른 검사가 될 수 있다는 점에 유의하여야 한다. 한 예로 WMS-R(1987)의 LM 소검사 두 번째 이야기인 '교통사고'를 주제로 한 이야기의 경우 트럭의 크기 및 부품에 대한 내용은 국내 피검자에게 매우 낯선 내용들이다. 미국에서는 이미 오래 전부터 자동차가 필수 생활 수단이기 때문에 일반인들이 자동차에 대한 지식이 풍부한 반면 우리 사회에서는 아직 자동차 문화 역사가 짧아 관련 지식을 갖고 있지 않은 사람들이 더

많다. 따라서 이 검사를 국내에서 실시하게 되면 이것은 기억 검사이기보다는 자동차에 대한 지식 혹은 미국 문화의 친숙성을 측정하는 검사가 되기 쉽다. 이 이야기의 경우 주제의 친숙성은 노인들 특히 여성 노인 인구의 경우 매우 떨어지게 되므로 노인 기억 평가 시 타당성이 더 떨어질 가능성이 높다.

한국 노인들은 WMS이 개발된 미국이나 영국에 비해 정규 교육을 받은 평균 연한이 짧고 편차가 크다(통계청, 1995; U.S Census Bureau, 1997). 이는 전반적인 신경심리 수행의 저하와 수행 범위의 증가로 이어진다(최진영, 이소애, 1997). 더더구나 검사 수행에 있어 모집단에서 그와 유사한 검사를 받은 인구가 많아질수록 수행이 향상되는데(Flynn, 1987) 인지 기능 검사나 신경심리 검사가 교육 기회가 적었던 국내 노인들에게는 생소하여 수행이 더 낮게 평가될 가능성도 있다. 이에 아직 한국 규준이 수집되지 못한 번역본의 원규준을 사용하게 되면 진단 및 판단 오류의 가능성이 커지며 특히, 잘못된 정적(false positive) 진단-제 1종 오류-를 범할 가능성이 높아진다.

현재 국내 노인 인구에 실시할 수 있는 표준화된 검사로는 한국형 치매 평가 검사 K-DRS(최진영, 1998) 내 기억 소검사가 있어 초기 DAT 변별에 유용하다고 보고되었다(최진영, 1998). 그러나 이 소검사는 기억 감퇴의 세부적인 특징들을 자세하게 파악하기에는 너무 단순하여 DAT 전 단계(preclinical phase)에서 일어나는 미세한 기억력 저하를 확인할 수 있는 좀 더 예민한 언어 기억 검사의 필요성이 제기되었다. 오래 전부터 서구에서는 여러 가지 이야기 회상 검사가 개발되고 이에 따른 규준 연구가 활발히 진행되어 왔으나(Rapaport, Gill & Schafer, 1968; Wechsler, 1997) 현재 국내에서는 Memory Assessment Scales(Williams, 1991)를 번역, 수정한 한국판 MAS(한국

판 기억 평가척도; Korean Memory Assessment Scale: K-MAS, 1999)의 이야기 회상 검사만이 있다. K-MAS의 이야기 회상검사는 자유회상 시 피검자가 회상하여 말한 그대로를 기록하여 채점하는 WMS 방식과는 달리 이야기에 관한 질문에 대하여 피검자가 말하는 단서 회상 형식으로 채점하게 되어 있어 이야기를 사건의 논리적 흐름에 맞게 구성하는 능력을 측정하기 어렵게 되어 있다. 또한 이 검사는 노인을 위하여 개발된 것이 아니어서 노인에게 보다 친숙한 내용의 이야기 검사의 개발이 필요하였다. 최근 고연령의 인구가 증가하고 있으며 특히, 무학을 포함한 6년 이하의 교육을 받은 고령 노인 인구가 상당히 많은 우리나라의 노인 집단의 특성을 고려할 때, 이들에게 보다 친숙한 주제의 이야기 회상 검사의 개발과 노인 집단을 대상으로 한 규준 연구는 중요하다.

신경 심리 검사를 개발하고 표준화 과정에는 그 검사의 목표 인구가 될 모집단에서 어떤 인구통계학적 요인이 고려되어야 하는지 판단하는 것이 중요한데, 이야기 회상 검사 수행에서도 유의미한 영향을 미치는 인구 통계학적인 변인들에 대하여 많은 논의가 있어 왔다(Abikoff, Alvir, Hong et al., 1987; Hulicka, 1966; Ivison, 1986; Lezak, 1995; Margolis & Scialfa, 1984; McCarty, Logue and Siegler, 1982). 이 연구 결과들을 요약하면 피검자가 연령이 많을수록, 그리고 교육을 적게 받았을수록 이야기 회상 검사에서의 수행이 떨어지는 경향이 일관되게 관찰되었다. 반면에, 성별 변인에 대해서는 연구에 따라 여자가 또는 남자가 더 좋은 수행을 보이는 것으로 관찰되거나(Verhoff, Albert, Kaplan, et al, 1979), 성별에 따른 차이가 없다는 등 불일치하였다. 이에 따라 국내 인구를 대상으로 한 이야기 회상 검사의 표준화 과정에서 어떤 인구 통계학적인 변

인들이 수행에 유의미한 영향을 미치는지 확인하는 것이 필요하다.

본 연구의 목적은 한국의 문화와 언어 특징을 반영하고 노인 인구에게 친숙한 이야기 회상 검사를 개발하여 검사의 신뢰도와 타당도를 확인하고 검사 수행에 영향을 미치는 인구 통계학적인 변인들을 고려하여 그 기준을 제시하고자 한다. 뉴스를 통해서 자주 접하게 되는 사건인 유괴를 주제로 국내 노인들이 쉽게 들어본 이름이나 지명을 사용하여 이야기를 구성하였다. 기준은 만 55세~89세 정상 장·노년층을 대상 인구로 삼아 치매 평가에 도움이 되도록 하였다. 흔히 제시되는 노인 기준에 비해 10년 더 젊은 인구의 기준도 마련하였는데 이는 치매의 발병이 50대부터 시작된다는 역학 연구들과(Ferraro & Jervis, 1941; Molsa, Marttila, & Rinne, 1982) 임상 경험에 근거하였다. 또한 이야기 회상 검사에서 유의미한 인구통계학적 요인에 따라 기준을 산출하였기 때문에 보다 정확한 해석을 가능할 수 있을 것으로 기대되었다.

검사 구성 및 절차는 최근 출판된 Wechsler Memory Scale-III(WMS-III, 1997)의 LM 소검사의 형식과 유사하게 즉각 회상, 지연 회상 및 재인 과제로 구성되었다. 이는 언어적 이해, 부호화 저장, 인출 등 기억 과정과 특히 이야기 회상 시 필요한 사고를 조직화하는 개념화 능력과 관리 기능을 측정할 수 있도록 하여 한국 정상 노인들의 기억 세부 특징들을 확인할 수 있었다. 또한 LM와 비교형을 개발하므로 영미권의 많은 연구에서 보고되는 Wechsler Memory Scale 사용 연구들과 비교 가능하게 하여 국내 연구의 국제화 및 해외 연구의 국내 해석에 도움이 될 것으로 생각되었다.

연구방법

연구 대상

연구 참여자들은 기준 마련을 위하여 일반 지역 사회에서 표집된 정상 노인 273명이었다. 세 개의 연령 집단, 55-64, 65-74, 75세 이상과 두 개의 학력 집단, 0-6년, 7년 이상으로 구분하여 총 6개 기준 집단을 비슷한 크기로 표집하였다. 이는 한국판 치매 평가 검사(Korean-Dementia Rating: K-DRS)의 기준 연구(최진영, 1998)에서 두 개의 학력 수준(6년 이하와 7년 이상)과 두 개의 연령 수준(만 55-64세와 만 65-84세)이 유의한 수행 차이를 보였다는 결과에 근거하였다. 다만, 연령 집단의 경우 노화됨에 따라 일어나는 인지적인 변화를 좀 더 예민하게 관찰하고 싶었기 때문에 3개의 집단으로 표집하였다.

표집은 두 번에 나누어서 시행되었는데 1차 표집에서는 136명이 서울과 부천, 의정부, 성남, 분당 등을 포함한 수도권 지역과 충남 당진에서 참여하였으며 2차 표집에서는 137명이 서울, 인천, 용인을 포함한 수도권 지역, 예산, 대구, 마산에서 그 지역에 위치한 종교 기관, 복지 기관 및 직장, 가정에서 참여하였는데 한 기관이나 가정에서 10명 이상 검사하지 않도록 하였다. 참여자 모두는 K-DRS에서 정상적인 인지 수행 능력을 보였고, 간단한 신경심리학적 인터뷰인 Christensen(1991)의 건강 선별 배제 기준(Health Screening Exclusion Criteria)에 부합하는 사람들 중에서 연령이 만 55세 이상 89세 이하인 자들이었다. 1, 2차 표집 집단의 K-DRS 수행 점수를 비교하였을 때 집단 간 K-DRS에서 통계적으로 유의미한 차이가 없어($F(1,271)=.005, p=.945$) 이 후 모든 통계 처리는 통합하여 처리하였다. 총 273명의 참여자들의 인구통계학적 특성을 표 1에

표 1. 정상 노인 참여자들의 인구 통계학적 특성

연령	교육	사례수	연령 ($M \pm SD$)	교육 ($M \pm SD$)	남:여
55~64세	0~6년	43	60.05 \pm (2.81)	4.54 \pm (2.60)	13:30
	7~9년	15	60.40 \pm (2.47)	9.00 \pm (.01)	6:9
	10~12년	17	60.06 \pm (2.77)	11.82 \pm (.53)	11:6
	13년이상	13	58.23 \pm (2.98)	15.85 \pm (1.86)	8:5
		88	59.84 \pm (2.81)	8.36 \pm (4.56)	38:50
65~74세	0~6년	48	70.12 \pm (2.65)	3.60 \pm (2.73)	7:41
	7~9년	18	70.17 \pm (3.40)	8.06 \pm (1.00)	5:13
	10~12년	15	70.80 \pm (3.00)	11.20 \pm (1.01)	4:11
	13년이상	16	69.81 \pm (2.69)	15.94 \pm (1.88)	11:5
		97	70.19 \pm (2.83)	7.64 \pm (5.12)	27:70
75~89세	0~6년	49	80.02 \pm (3.52)	2.16 \pm (2.47)	10:39
	7~9년	12	78.58 \pm (2.47)	8.00 \pm (.85)	2:10
	10~12년	19	78.58 \pm (3.36)	10.63 \pm (.83)	4:15
	13년이상	8	78.63 \pm (5.01)	14.38 \pm (4.95)	2:6
		88	79.39 \pm (3.53)	5.90 \pm (4.87)	18:70
전	체	273	69.82 \pm (8.44)	7.31 \pm (4.95)	83:190

정리하여 제시하였다. 또한 채점자간 신뢰도를 구하기 위하여 본 연구에 참여한 273명의 정상 노인을 대상으로 표집 시 고려한 연령(55~64세, 65~74세, 75~89세) 및 교육(6년 이하, 7년 이상)의 각 하위 집단(표 1 참고)에서 5명씩, 총 30명을 선발하였다.

검사 도구

이야기 회상 검사(Story Recall Test; 이하 SRT)

이 검사는 이야기를 들려준 직후와 20-30분의 시간이 경과한 후 회상하도록 하는 검사이다. 보통 이야기 회상 패러다임에서 이야기는 이해하고 기억하기 쉬운 내용으로 구성하는 것을 원칙

으로 하여 사고나 사건에 대한 기술을 논리적인 순서로 제시하게 되어있다(Lezak, 1995). 이 검사는 한국 사회에 친숙한 사건으로 유괴사건을 주제로 한 이야기를 다루고 있는데 형식 구성은 24개의 이야기 단위(의미 있는 어절)들로 구성되었다. 검사의 실시는 개인 검사로 진행되었으며 이야기를 청각적으로 들려주고 이야기에 대하여 즉각적 자유회상, 30분 후 지연 자유회상이 요구되었고 마지막으로 이야기에 관한 10문항의 재인 검사가 주어졌다.

채점 기준은 의미를 중시하여 불러준 말 그대로의(literal) 회상과 함께 동일한 의미의 대치 반응들에 대해서 점수를 주도록 하였다. 이야기 회상 과정에서 흔히 피험자들은 들려 준 그대로를

말하기보다 다양한 대치 반응들을 보이기 때문에 채점 기준에 대한 논란이 많았으나 의미 중심적인 채점이 기억력 장애를 변별하는데 더 타당한 것으로 밝혀져(Wechsler, 1987) 의미의 왜곡 없이 회상할 수 있으면 점수를 부여하는 것이 추세이다. 따라서 본 연구의 이야기 회상 검사의 채점 기준은 불러 준 그대로를 정확하게 반복한 것에 1점을 주는 것을 원칙으로 하면서 이와 더불어 의미상 본래의 의미를 상실하지 않은 동의 어휘를 사용한 경우에도 1점을 주는 것으로 하였다. 채점에 있어서 얼마나 정확한 것까지 점수를 부여할 것인지에 대한 기준을 명확히 하기 위해서 각 이야기 단위 1점에 해당하는 예와 0점에 해당하는 예를 자세하게 기술하여 참고하도록 하였다. 또한 각 이야기 단위의 내용 중 일부를 생략한 경우라든지, 해당 이야기 단위 기준에는 충분치 않으나 이야기 전개상 논리적으로 맞고 의미가 적절한 어휘를 사용한 경우에는 0.5점을 주었다. 기존 연구에서 부분 점수를 인정해주는 채점 방식이 더 신뢰롭고 타당성이 있다는 보고가 있고(Abikoff, et al., 1987) 또 저학력자가 많은 우리 노인 인구의 특성을 감안할 때 이 채점 체계가 좀 더 타당성이 있다고 생각되었다. 또한 이야기 단위에 더하여 이야기의 주제 단위도 측정하였는데 주제 단위란 논리적 흐름상 가장 중요한 의미를 지닌 주제 내용으로 이는 이야기 단위 보다 일반적인 정보에 대한 회상 능력을 측정하기 위함이다(Wechsler, 1997). 채점 기준은 노인 수행의 경험적인 자료를 근거로 신경심리학자 1명과 3명의 석사급 대학원생들로 구성된 평가팀이 채점상의 의문사항 및 의견을 나누는 회의를 통하여 결정되었다. 이야기 회상 검사의 측정치는 전체 단위 수인 24점과 이야기 6개의 주제 단위수를 더한 30점을 각 각 총점으로 하는 즉각적 자유회상 점수와 지연 자유회상

점수 그리고 10문항으로 구성된 10점 만점의재인 점수가 있다. 이외에도 기억의 파지 (retention) 능력을 측정하는 보유율인 (Saving Score)를 계산하였는데 이는 즉각적으로 회상한 내용을 지연 후에 얼마나 보유하고 있는지를 지수화한 것으로 $[(\text{지연 자유회상 점수}/\text{즉각적 자유회상 점수}) \times 100]$ 공식을 이용하였다.

한국판 치매 평가 검사(Korean-Dementia Rating Scale; K-DRS)

노인들의 전반적인 인지 능력을 측정하기 위하여 사용되는 한국형 치매 평가 검사(K-DRS; 최진영, 1998)는 주의(Attention), 관리기능(Initiation & Perseveration), 구성(Construction), 개념화(Conceptualization), 기억(Memory)의 세부 인지 능력들을 측정하는 5개의 소검사들로 구성되었다. 주의 37점, 관리기능 37점, 구성 6점, 개념화 39점, 기억 25점을 만점으로 하며, 총점은 144점이다. 본 연구에서는 참여자의 정상적인 인지 상태를 확인하고 이야기 회상 검사의 구성 타당도를 확인하기 위하여 실시되었다.

노인용 언어 학습 검사(Elderly Verbal Learning Test: 이하 EVLT)

노인용 언어 학습 검사(김명진, 2002)는 언어적 기억력, 학습 능력 및 간섭 취약성을 측정한다. 단어 목록은 세 가지 의미 범주에서 세 단어씩 모두 9 단어로 구성되었다. 단어들은 교육과 성별, 계층의 영향이 적은 단어들로 선택되었다. 학습 목록의 의미 범주는 과일, 동물, 가전 제품 3가지였으며 간섭 목록의 9단어 중 3개의 단어는 학습 목록과 같은 범주(“동물”) 같았다. 재인 목록은 총 30단어였다. 검사 절차는 우선 학습 목록을 다섯번 불러주고 매회 회상하도록 하는 5회의 학습 시행 후 1회의 간섭 목록 즉각 회상

을 실시한 후, 학습 목록의 단기 자유 회상과 단기 단서 회상을 실시하였다. 20-30분 동안의 지연 후, 지연 자유 회상과 지연 단서 회상을 실시하고 마지막으로 학습 목록의 재인 검사를 실시하였다. EVLT의 측정치들로는 5회의 시행에 걸쳐 옳게 회상한 단어 수의 총합인 학습 총량, 간섭 시행, 단기 자유 회상 시행, 단기 단서 회상 시행, 지연 자유 회상 시행, 지연 단서 회상 시행, 재인 시행에서 학습 목록 단어를 긍정 재인한 개수와 학습 목록 단어 이외의 단어를 부정 재인한 개수를 합한 재인 회상의 6가지였다. 각 회상 시행은 9점이 최대치이며, 학습 총 회상은 45점, 재인 점수는 30점이 최대치였다. 이 검사는 본 검사 바로 전에 실시되었으며 본 검사와의 상관 분석을 통하여 구성 타당도를 확인하기 위해 실시되었다.

연구 절차

본 검사인 이야기 회상 검사의 실시가 있기 전에 모든 정상 참여자들에게 간단한 신경심리학 인터뷰 실시를 통하여 사전 선별 절차가 진행되었다. 이 과정은 참여자의 기본적인 인적 자료를 수집하고 참여자가 연구 참여자의 참여 기준(criteria of inclusion)에 부합되는지 여부를 검토하려는 목적으로 이루어졌다. 이러한 선별 과정을 거친 후 K-DRS를 실시하여 연구 참여자들의 전반적인 인지 능력을 측정하기 위해 실시되었다. 이어서 신경 심리 검사로 노인용 언어 학습 검사의 학습 목록 즉각 회상 5회, 간섭 목록 즉각 1회, 학습 목록 단기 자유 회상, 단기 단서 회상이 실시되었다. 그 다음에 본 검사인 노인용 이야기 회상 검사를 청각적으로 들려주고 그 이야기에 대한 즉각적 자유 회상이 실시되었다. 즉각 회상과 지연 회상 사이의 20-30분 동안에는

언어 기억에 영향을 미치지 않도록 비언어적 간섭과제로는 숫자 폭 바로 따라하기 검사와 거꾸로 따라하기 검사(송호정, 2002), 그리고 시공간 폭 따라하기 검사와 거꾸로 따라하기 검사(송호정, 2002)와 단순화된 Rey 도형 검사(김은영, 2001)가 표준화된 절차에 따라 주어졌다. 이어서 노인용 언어 학습 검사의 장기 자유회상, 장기 단서 회상, 재인이 실시되었으며 그 직후 본 검사인 이야기 회상의 지연 회상과 그리고 이야기에 관한 10문항의 재인 회상이 주어졌다. 검사 시 노화로 인한 시력이나 청력의 약화와 같은 비인지적인 기능 장애가 검사 수행에 영향 미치지 않도록 초기 인터뷰에서 이를 확인하였고, 필요하다면 돋보기를 사용할 수 있도록 제공하거나 지시문을 크게 읽어 주었다. 만일 이러한 도움에도 검사가 불가능할 경우에는 검사 실시를 중단하였다. 정상 집단의 검사는 훈련받은 임상 심리학 대학원생 4명이 실시하였다.

자료 분석

통계적 검증은 SPSS for Windows(Version 11.0)를 이용하였으며 분석은 다음과 같다. 첫째, 채점자 간 신뢰도를 알아보기 위해 Pearson 상관 분석을 실시하였으며, 기억 검사로서의 구성 타당도를 확인하기 위해서 이야기 회상검사의 세가지 측정치들과 K-DRS의 점수들, EVLT 검사의 수행치들과 부분 상관계수를 교육, 연령, 성별이 과제 수행에 미치는 영향을 통제한 후 구하였다. 둘째, 인구 통계학적 변인들과 이야기 회상 검사의 수행간의 관계를 알아보기 위해서 분석할 변인을 제외한 나머지 변인들은 통제하는 식으로 각 인구 통계학적 변인들에 대해 차례로 부분 상관 분석을 실시하고, 교육, 연령, 성별을 독립변인으로 하고 즉각적 자유회상, 지연 자유회상, 재인

회상을 각각 종속변인으로 하여 단계식 중다 회귀 분석(step-wise multiple regression analysis)을 실시하여 이야기 회상 검사 세 가지 측정치들 각각에 영향을 미치는 인구 통계학적인 변인들을 확인하였다. 셋째, 교육 집단 간, 연령 집단 간, 그리고 성별 간 이야기 회상 검사 수행치들에서 차이가 있는지를 확인하기 위하여 공변량 분석을 실시하였다. 본 연구에서 관찰된 인구 통계학적 변인들간의 상관 결과에서 노인들의 연령이 증가함에 따라 교육수준은 낮아지고 여자 노인의 교육 수준이 낮다는 것이 드러나고 있어 인구 통계학적 변인들간의 상호작용을 통제할 필요가 있어 변량 분석 대신 공변량 분석을 실시하였다. 그 후 기준 집단을 세분화하는 결정을 하기 위해 사후 검정을 실시하였는데 인구통계학적 집단에 따른 세부적인 기준을 작성하기 위하여 집단 간 차이를 가장 예민하게 확인하는 Duncan 검증을 실시하였다(Howell, 1987). 넷째, 확인된 인구통계학적 변인에 따라 기준 집단을 세분화하여 각 집단의 기술 통계치로 평균과 표준 편차를 기술하여 제시하였다. 마지막으로 정상노인들의 언어적 기억수행 특징을 알아보고자 이야기 회상 수행치들 간의 Paired t-test를 실시하였다.

결 과

채점자간 신뢰도와 구성 타당도

2명의 채점자간 신뢰도는 즉각적 자유 회상에서 $r(273) = .98$ ($p < .001$), 지연 자유회상에서 $r(273) = .99$ ($p < .001$)로 매우 높은 채점자간 일치율을 보였다.

구성 타당도를 확인하기 위하여 교육, 연령, 성별이 과제 수행에 미치는 영향을 통제한 후 본 검사인 이야기 회상검사의 세 가지 수행치들과 K-DRS의 총점 및 5개의 소검사 점수들, EVLT 검사에서 5회 반복학습을 통해 습득된 학습 총점, 단기자유회상, 장기 자유회상, 재인 회상 점수간의 부분 상관계수를 구하였다. 그 결과, K-DRS 총점과 기억 소검사와는 매우 유의미한 상관을 보였으며 ($p < .001$), 그 외 소검사 관리기능, 개념화, 주의력과도 유의미한 상관을 보였다 ($p < .01$). 그러나 구성 소검사와는 상관이 없었다. 노인용 언어 학습 검사의 수행치들과는 모두 유의미한 상관을 보였는데 학습 총 회상 점수와 단기 자유회상, 장기 자유회상, 재인 회상과 상관이 높았다 ($p < .001$). 이를 표 2에 제시하였다.

표 2. 이야기 회상 검사와 다른 검사와의 부분 상관 계수

	1	2	3	4	5	6	7	8	9	10
즉각적 자유회상	.322 **	.166 *	.200 *	.021	.196 *	.324 **	.303 **	.219 **	.286 **	.268 **
지연 자유회상	.299 **	.162 *	.176 *	-.023	.178 *	.329 **	.313 **	.294 **	.309 **	.273 **
재인 회상	.260 **	.167 *	.160 *	.014	.166 *	.234 **	.229 **	.223 **	.257 **	.231 **

1=K-DRS 총점, 2=K-DRS 주의, 3=K-DRS 관리기능, 4=K-DRS 구성능력 5=K-DRS 개념화 능력
6=K-DRS 기억, 7=EVLT 5회 반복 시행 후 학습 총 회상점수, 8=EVLT 단기 자유회상,
9=EVLT 장기 자유회상 10=EVLT 재인회상

* $p < .01$, ** $p < .001$

표 3. 인구통계학적 변인들과 이야기 회상 검사의 측정치들과의 부분 상관 계수

	즉각적 자유회상	지연 자유회상	재인회상
교육	.355 ***	.321 ***	.386 ***
연령	-.010	-.156*	-.193**
성별	.155 *	.122 *	.138 *

* $p < .05$, ** $p < .01$, *** $p < .001$

인구통계학적 변인들과 이야기 회상 검사와의 상관 관계

인구 통계학적 변인들과 이야기 회상 검사 측정치들간의 상관 관계를 알아보는데 인구 통계학적 변인들 간의 상관관계가 유의미하게 나타나고 있어(나이와 교육간의 상관관계수는 $r = -.220$, $p < .001$, 나이와 성별간의 상관관계수는 $r = .147$, $p < .05$, 교육과 성별과의 상관관계수는 $r = -.324$, $p < .001$), 관심 변인 외의 효과를 통제하는 부분 상관 분석을 실시하였다.

유목 변인인 성별과 이야기 회상 검사의 수행치들 간의 상관 분석은 point biserial correlation을 실시하고 이를 표 3에 제시하였다. 그 결과, 교

육 수준이 높을수록 수행의 향상을 보이는데 반해 연령이 증가할수록 수행의 저하를 보이며 남자보다 여자가 더 나은 수행을 하는 것으로 나타났다(남자를 1로 여자를 2로 코딩하였음).

이야기 회상 검사에 미치는 인구통계학적인 변인들의 영향력

인구 통계학적인 변인들이 이야기 회상 검사 수행에 얼마나 영향을 미치는지 알아보기 위해 교육, 연령, 성별을 독립변인으로 하고 즉각적 자유회상, 지연 자유회상, 재인 회상을 각각 종속변인으로 하여 단계식 중다 회귀 분석을 실시하였다.

표 4. 인구통계학적 변인의 다중 회귀 분석 결과 요약

종속변인	예언변인	R	R ²	R ² change	Beta	F change
즉각적 자유회상	교육	.348	.121	.121	.395***	37.334***
	성별	.374	.140	.019	.145*	5.947*
지연 자유회상	교육	.332	.110	.110	.337***	33.482***
	연령	.359	.129	.019	-.151*	5.869*
	성별	.377	.142	.013	.120*	4.033*
재인 회상	교육	.399	.159	.159	.401***	51.233***
	연령	.432	.187	.028	-.181**	9.173**
	성별	.449	.202	.015	.131*	5.181*

* $p < .05$, ** $p < .01$, *** $p < .001$

그 결과, 즉각적 자유회상에서는 교육과 성별이 12~14%의 설명력을 지니고 있는 것으로 나타났고, 지연 자유회상에서는 교육, 연령, 성별이 약 11~14%의 설명력을, 재인 회상에서는 교육, 연령, 성별이 약 16~20%의 설명력을 지니고 있는 것으로 나타났다. 여기에서 예언변인들 중 교육 변인의 beta값이 모든 측정치에서 가장 높게 나타내고 있어 교육이 가장 큰 영향을 미치는 인구 통계학적 변인임이 확인되었다. 자세한 단계식 중다 회귀 분석의 결과는 표 4에 제시하였다.

인구 통계학적 변인들에 따른 이야기 회상 검사의 수행 비교

교육 집단 간(0~6년, 7~9년, 10~12년, 13년 이상) 수행 차이를 알아보기 위하여 연령과 성별을 공변인으로 하여 이야기 검사 수행치들을 공변량 분석하였다. 교육 집단별 차이가 즉각적 자유회상에서는 $F(3, 267)=12.49, p<.001$, 지연 회상에서는 $F(3, 267)=10.81, p<.001$, 재인 회상에서는 $F(3, 267)=14.16, p<.001$ 로 교육수준에 따른 이야기 회상 검사 세 가지 측정치들이 모두 유의미한 차이를 보이는 것으로 나타났다. 보다 타당한 기준을 작성하기 위하여 교육, 연령 및 성별에 따른 집단간 차이를 Duncan 검증(Howell, 1987)을 $\alpha=0.05$ 수준에서 실시하였다. 그 결과 교육 집단에서는 재인 검사에서만 0~6년과 7~9년을 하나의 하위조(subset)로 또 10~12년과 13년 이상을 다른 하위조로 나뉘었다.

연령의 경우도 세 집단(55~64세, 65~74세, 75~89세) 간에 검사 수행 차이가 있는지를 알아보기 위하여 교육과 성별을 공변인으로 하여 공변량 분석을 하였다. 그 결과, 연령 집단별 차이가 즉각적 자유회상에서는 유의미한 차이를 보

이지 않았으나 $F(2, 268)=1.69, ns$. 지연 회상에서는 $F(2, 268)=1.93, p<.05$, 재인 회상에서는 $F(2, 268)=1.75, p<.05$ 로 유의미한 차이가 관찰되었다. 이에 지연 자유회상과 재인 회상 수행치들의 연령 집단차이를 알아보기 위해 위와 동일한 방법으로 Duncan 사후검증을 한 결과 지연회상과 재인회상 모두에서 74세 이하와 75세 이상의 두 집단으로 나누어지는 결과를 보였다.

동일한 방법으로 성별 간 이야기 회상 검사의 수행 차이를 보이는지를 알아보기 위하여 교육과 연령을 공변인으로 하여 공변량 분석을 실시하였다. 그 결과, 즉각적 자유회상에서는 $F(1, 271)=6.62, p<.05$, 지연 회상에서는 $F(1, 271)=4.03, p<.05$, 재인 회상에서는 $F(1, 271)=5.18, p<.05$ 로 성별에 따른 이야기 회상 검사 세 가지 측정치들 모두가 유의미한 차이를 보이는 것으로 나타났다. 이러한 결과를 토대로 이야기 회상 검사 세 가지 측정치 모두에서 성별에 따른 기준을 제시하기로 결정하였다.

이야기 회상 검사 수행의 성별, 연령 및 학력에 따른 기준

앞서 확인된 인구통계학적 변인에 기초하고 기준 집단의 일관성을 고려하여 기준 집단을 결정하였다. 교육의 경우 공변량 분석 결과 모든 수행치에서 교육 집단 별 차이가 매우 유의미한 것으로 나타났으며($p<.001$), 사후 검증 시 재인 수행에서 9년을 기점으로 두 집단으로 명확히 구분되어 이야기 회상 검사 세가지 수행치들에 대한 교육 기준을 9년을 기점으로 두 집단으로 나누는 것으로 통일하였다.

연령은 즉각적 자유회상을 제외하고는 공변량 분석에서 유의미하였고($p<.05$), 사후 검증에서 75세를 기준으로 집단이 나누어져 지연 자유회상

표 5. 학력 수준과 성별에 따른 즉각적 자유회상의 평균과 표준편차

측 정 치	교 육	성 별	사례수	평 균	표준편차
즉각적 자유회상	≤9	남	43	13.29	5.25
		여	142	13.88	5.48
	〉 9	남	40	16.34	5.25
		여	48	18.17	4.13
전	체		273	14.90	5.46

과 재인회상 두 수행치의 규준에서만 2개의 연령 집단을 구성하였다.

성별에 있어서도 모든 수행치에서 유의미한 차이를 보였기에($p < .05$) 성별 규준을 작성하였다.

구체적으로 즉각적 자유회상은 0~9년과 10년 이상인 두 학력 집단과 그에 따른 남·여의 두 성별 집단으로 나누어 평균과 표준편차를 표 5에 기술하여 규준을 제시하였다.

그리고 지연 자유회상과 재인 회상은 연령을 55~74세와 75세~89세의 두 집단으로 나누고 다시 학력을 0~9년과 10년 이상인 두 개 집단으로 나눈 뒤 성별로 세분화 한 총 8집단의 지

연 자유회상과 재인 회상의 평균과 표준편차를 표 6, 표 7에 기술하여 규준을 제시하였다.

각 규준 집단에 대하여 이야기 회상 검사의 측정치들이 정규분포를 따르고 있는지를 검증하기 위해 단일 표본 Kolmogorov-Smirnov 검증 결과를 살펴 본 바에 의하면 즉각적 자유회상과 지연 자유회상 수행치들은 정상 분포를 보이고 있는데 반해 재인 회상의 경우는 왜도 = -.424로 왼쪽으로 편포된 분포를 보였다. 따라서 재인 회상의 경우 중앙치와 왜도를 같이 제시하였다.

표 6. 연령과 학력 수준, 성별에 따른 지연 자유회상의 평균과 표준편차

측정치	교육	연령	성별	사례수	평균	표준편차
지연 자유회상	≤9	55~74세	남	31	12.26	4.83
			여	93	13.07	5.46
		75~89세	남	12	11.92	5.63
			여	49	10.55	4.90
	〉 9	55~74세	남	34	15.35	6.07
			여	27	17.24	4.49
		75~89세	남	6	11.75	4.79
			여	21	15.71	4.21
전		체	273	13.34	5.63	

표 7. 연령과 학력 수준, 성별에 따른 재인 회상의 평균과 표준편차

측정치	교육	연령	성별	사례수	평균	표준편차	중앙치	왜도
재인회상	≤9	55~74세	남	31	6.84	1.39	7.00	-.405
			여	93	7.43	1.59	8.00	-.441
		75~89세	남	12	7.33	1.37	7.50	-.217
			여	49	6.39	1.64	6.00	-.042
	> 9	55~74세	남	34	8.27	1.48	8.27	-.723
			여	27	8.67	1.14	9.00	-.611
		75~89세	남	6	7.00	1.10	7.00	-1.37
			여	21	8.14	1.46	9.00	-1.23
전 체				273	7.44	1.65	8.00	-.424

정상 노인 집단의 기억 수행의 특징들

이야기 회상 검사의 여러 측정치들로 정상 노인들의 다양한 학습 및 기억수행 특징들을 살펴 보았다. 먼저 자유회상과 단서회상 수행간의 유의미한 차이가 있는지를 알아보기 위해 즉각적 자유회상과 재인 회상, 지연 자유회상과 재인 회상과의 Paired T-test를 실시하였다.

그 결과, 즉각적 자유회상과 재인 회상 수행간의 차이($t=-24.24$, $p<.001$)와 지연 자유회상과 재인 회상 수행간의 차이($t=-28.48$, $p<.001$) 두 가지 모두에서 유의미한 차이가 나타났다. 즉각적 자유회상과 지연 자유회상 간을 비교하기 위해 Paired T-test를 실시한 결과, 즉각적 자유회상과 지연 자유회상간의 유의미한 차이($t=9.773$, $p<.001$)가 관찰되었다. 또한 정상 노인들이 즉각적 회상에서 획득한 정보를 20~30분 정도의 지연을 두고 다시 회상할 시 보유하고 있는 학습의 양을 측정하기 위하여 보유율 또는 파지율(saving or retention score)을 구하였다. 보유율은 즉각적 이야기 단위 자유회상의 수행정도에 영향을 받

지 않기 위하여 $[(\text{즉각적 자유회상 점수} - \text{지연 자유회상 점수}) \times 100]$ 으로 계산되었다.

본 연구의 정상 노인 집단 273 명의 평균 보유율은 90.63%로 즉각적 자유회상 시 회상한 내용의 약 91%를 지연 자유회상 시 다시 회상한다는 결과를 보여 주었으며 보유율의 수행 범위는 31-200%이었다. 273명 중 68명인 25%는 즉각적 자유회상 보다 지연 자유회상 시 더 나은 회상을 보여 주었다.

논 의

본 연구는 한국 노인 인구에 실시할 수 있는 노인용 이야기 회상 검사를 개발하여 신뢰도와 타당도를 확인하고 이 검사 수행에 미치는 인구 통계학적 변인을 검토하여 이에 기초한 기준을 마련하고자 하였다. 기억 검사로서의 타당도를 검토하기 위해서 K-DRS의 총점 및 소검시들과 이야기 회상 검사의 세 가지 측정치들과의 상관 분석을 실시하였다. 이야기 회상 검사의 모든 수

행치들은 기억 외에도 주의, 개념화, 관리기능 소검사들과 유의미한 상관을 보였다($p < .01$). 이는 이야기 회상 검사가 기억과 함께 언어 정보를 이해하고 정리하여 개념화하는 기능들을 평가하는 검사임을 보여주는 결과로 일상 생활에서의 기억 기능을 측정하는 검사로서의 구성 타당도를 지지하는 결과라고 할 수 있다. 한편 언어적 기억과는 개념적으로 연관이 적은 구성 능력을 측정하는 '구성' 소검사와는 상관이 없는 것으로 관찰되어 변별 타당도를 확인할 수 있었다. 순수한 언어 기억 및 학습 능력을 가장 잘 측정한다는 언어 학습 검사인 EVLT 수행치들과 일관적으로 높은 상관이($p < .001$) 관찰된 것은 언어 기억 검사로서의 이야기 회상 검사의 수렴 타당도를 지지해 준다. 검사의 신뢰도를 확인하기 위하여 채점의 준거와 그에 해당하는 예를 마련하고 두 명의 채점자간 상관을 검토하여 본 결과, 채점자들 사이의 높은 일치율을 보였다.

연령, 학력, 및 성별에 따른 기준을 개발하기 위하여 인구통계학적 변인들과 이야기 회상 검사 수행과의 상관 및 회귀 분석을 실시한 결과 세 변인 모두 영향을 미치고 있는 것이 밝혀졌다. 특히 정규 교육 연한이 세 수행치 모두에 큰 영향력을 미쳤는데($p < .001$) 교육 연한이 짧을수록 수행이 감소하였다. 이 검사에서는 들려준 이야기를 시간과 사건의 맥락에 따라 재조직화하여 이해하는 능력을 필요로 하는데 이러한 언어적 개념화는 교육과 밀접한 관련이 있는 것으로 밝혀지고 있다(최진영, 나덕력, 박선희, 1998). 연령 효과는 지연과 재인 회상 수행에서 유의미했는데 이는 언어적 학습 및 기억 능력이 연령 증가에 따라 감퇴된다는 다른 연구 결과들과 일관된다(Craik, 1990; Kaszniak, Poon & Riege, 1986). 교육과 연령의 효과는 대부분의 신경심리 검사의 기준 연구들에서 이들 요인이 검사 수행과

밀접한 상관이 있다는 연구 결과들과도 일치된다(김정기 & 강연옥, 1998; Crossen & Wiens, 1994; Lezak, 1995).

성별은 모든 수행에서 영향을 미치고 있어 이를 반영하여 기준을 산출하였는데 여성 노인들의 수행이 더 좋았다. 이것은 대부분의 인지 기능 수행에서 여자 노인들의 수행이 낮게 관찰되고 이때 교육의 효과를 제거하면 성별 간 차이가 없어지는 것과는 대조를 이룬다. 좀 더 자세히 설명하자면 이야기 회상 검사 수행치들에서 인구 통계학적 변인들이 모두 유의미하게 상관이 있었는데 이는 고연령일수록 저학력의 노인이 많았고 연령이 증가할수록 여자의 수가 더 많았으며 교육에 있어서도 여자가 남자에 비해 학력 수준이 낮았다. 여성 노인들이 나이가 많고 교육 수준이 낮음에도 이야기 회상 검사에서는 남자 노인들 보다 좋은 수행이 관찰되었다. 이것은 언어 기억 검사에서 자주 관찰되는 현상으로(Bleecker, et al., 1988; Bolla-Wilson, et al., 1986, Ivison, 1977) 대화를 통하여 인간 관계를 유지해 가는 여성 문화(Tremmel, 2003)와 관련되는 것으로 사료된다. 특히, 어떤 사건 이야기를 듣고 남에게 들은 대로 회상하는 이야기 회상 검사의 형태가 일상 생활에서 남자들보다 여자들에게 더 친숙하여 수행이 좋게 관찰된다는 주장이 있다(Ivison, 1986).

교육이 검사 수행에 중요한 요인으로 확인된 만큼 한국 노인의 이야기 회상 수행을 미국 노인의 것과 직접 비교하는 것은 무리가 있다. 개념화를 요구하는 신경심리 검사의 경우 저학력 인구에서 수행 저하가 두드러진다는 것을 감안할 때(Chey et al., 1999) 교육에 따른 기준의 개발은 한국 노인에게 필수적이다. WMS-R을 예로 보면 기준은 최소 12년의 학력자들을 대상으로 작성되어 있고 연령별 기준만을 제시하고 있는데,

60-69세 노인들은 즉각적 자유회상에서 평균 20.5(SD=6.4), 70-79세의 노인들은 평균=19.1(SD=6.7)의 수행이 관찰되었다(Abikoff, et al, 1987). 한편 본 연구의 이야기 회상 검사 즉각적 자유 회상에서는 65세 이상의 노인의 경우는 평균 수행이 14.5(SD=5.7)였지만 9년 이상의 학력자들의 경우 남자는 16.3(SD=5.3) 여자는 18.2(SD=4.13)였다. 신경심리 검사나 유사 심리 검사를 받아 본 노인(Flynn, 1987)의 비율이 미국 보다 훨씬 적다는 것과 언어와 이야기 내용이 달랐다는 것을 감안하면 미국 노인과 큰 차이가 없다고 할 수 있다.

이야기 회상 검사의 여러 수행치와 지수들을 이용하여 정상 노인들의 다양한 학습 및 기억수행 특징들을 살펴 본 결과, 먼저 자유회상보다 단서회상인 재인 회상에서 정상 노인 집단의 수행 향상이 관찰되었다. 이는 정상 노인들의 단서이용의 효율성을 입증해 주는 결과로 정상 노인 집단은 DAT 환자 집단과는 달리 단서가 제공될 때 기억 능력이 훨씬 향상된다는 선행 연구 결과들과 일치된다(Lezak, 1995; Spreen & Strauss, 1997). 인출 능력을 최대한 요구하는 자유회상보다 단서가 주어진 재인 회상이 정상 노인들에게는 더 쉬운 과제인 것이며 이는 인출 기능 저하를 시사하는 결과이기도 하다. 정상 노인들이 부호화나 저장 과정에 비해 인출 과정에서 취약한 것을 의미한다.

정상 노인 집단의 평균 보유율은 90.63%로 즉각 이야기 단위 자유 회상 시 회상한 내용의 약 91%를 지연 자유 회상 시 다시 회상한다는 결과이다. 흥미로운 사실은 전체 피험자 273명 중 68명인 약 25%는 즉각적 이야기 단위 회상 보다 더 나은 회상을 하였다. 이야기를 듣고 즉각적으로 회상하는 것이 다소 생소해 당황하는 경우도 있어 즉각 회상 수행의 저하가 있었을 수 있었

고 즉각적 자유 회상 과정을 통해 시연(rehearsal)할 수 있는 기회를 얻어 이야기를 조직화하는데 도움을 받아 수행이 향상되었을 수도 있다(Lezak, 1995). 이것은 초기 DAT 환자들이 짧은 지연 후(2-10분)에도 정보를 보유하지 못하고 급속한 망각을 보이는 것과 대조적이다(Welsh, et al., 1992). 본 연구에서 이야기 회상 검사 수행에서 관찰된 정상 노인 집단의 언어적 기억 특징들은 대부분의 언어 기억 검사의 선행 연구들에서 드러나고 있는 정상인들의 수행 특징들과 일치한다(Brandt, 1991).

본 연구의 제한점들을 언급할 필요가 있는데 언어 기억 검사로서 이야기 회상 검사로의 타당화를 보다 정교화하는 작업이 추후 연구에서 이루어져야 할 것이다. 언어적 학습 및 기억 능력과 관리기능의 손상이 두드러지게 나타나는 초기 DAT 환자들이나 기억 상실증 환자들에 대한 이야기 회상 검사의 임상적인 유용성을 살펴봐야 한다. 언어 학습 검사에 대한 선행 연구들은(Brandt, 1991; Kaszniak, 1986) 정상 노인 집단과 초기 DAT 환자 집단 기억 상실증 환자 집단 여러 종류의 치매 집단들이 검사 수행에서 차이를 보이고 있다고 보고하고 있어 이러한 변별 진단에 기여할 수 있는 임상적 유용성과 활용가능성을 높이도록 해야 할 것이다. 한편 서구권에서 치매 감별에 가장 타당도가 높다고 인정된 검사들이 한국 노인에게 사용되면서 서구에서는 관찰되지 않았던 교육 효과가 관찰될 경우 진단 도구로서의 타당도가 감소될 수 있어(Chey et al., 1998) 이에 대한 연구도 같이 병행되어야 할 것으로 사료된다. 국내 규준을 개발할 때는 교육 수준에 따라 세부적인 규준을 제시하는 것이 타당성 유지에 기여할 것이나 이외에도 다른 방법론적 보완을 모색하는 것이 필요할 수도 있다.

또한 본 연구는 규준 연구로 나름대로 국내

여러 지역의 55와 89세 사이 정상 노인들을 표집하였으나 이것이 전국적 층화 표집을 할 수 있으면 보다 더 신뢰로운 기준 확립에 도움이 될 것으로 생각되며, 세부 기준 집단의 크기가 고르지 않고 사례수가 적은 집단도 있어 보다 크고 균등한 크기의 기준 집단을 확보하는 추후 연구가 꼭 필요할 것으로 사료된다. 다만, WMS-R의 경우 16-74세 사이의 300명을 표본 대상으로 하였으며 55-64세, 65-69세, 70-74세 3개의 장노년 집단에 약 50명을 대상으로 3개의 연령별 기준만을 산출하였다. 이에 비해 노인용 이야기 회상 검사는 최근 증가하는 75세 이상 노인들도 표집하였으며 WMS-R의 두 배에 가까운 노인들을 대상으로 기준을 확립하였다.

참고문헌

- 김명진 (2002). 노인용 언어 학습 검사의 표준화 연구, 서울대학교 대학원 석사학위 청구논문.
- 김은영 (2001). 단순 Rey 도형 검사의 표준화 및 타당화 연구(노인 집단을 중심으로, 성신여자 대학교 대학원 석사학위 청구논문).
- 김정기, 강연옥 (1998). 한국판 캘리포니아 언어 학습검사(K-CVLT)의 표준화 연구. 한국심리학회지: 임상, 16(2), 379-395.
- 박은희, 최진영 (2000). 한국판 보스톤 이름대기 검사(K-BNT)의 단축형에 대한 노인 기준연구. 한국인지과학회 논문지, 11(2), 59-68.
- 보건복지부 (1998). 보건사회백서. 서울: 저자.
- 송호정 (2002). 한국 노인의 숫자폭 및 시공간폭 검사 표준화 연구, 서울대학교 대학원 석사학위 청구논문.
- 이현수, 안창일 (1999). 한국판 기억 평가 척도(K-MAS)의 타당도 연구, 고려대학교 대학원 박사학위 청구 논문.
- 최진영, 이소애 (1997). 한국판 치매 평가 검사(K-DRS)의 기준 연구, 한국심리학회지: 임상, 16(2), 423-433.
- 최진영 (1998). 한국판 치매 평가 평가: Korean-Dementia Rating Scale(K-DRS). 서울: 학지사.
- 최진영, 나덕렬, 박선희 (1998). 한국형 치매 평가의 타당도 연구, 한국심리학회지: 임상, 17(2), 140-147.
- 통계청 (1995). 인구주택 총조사 보고서. 서울: 저자.
- 통계청 (1998). 그래프로 보는 자료 [전자자료]. 서울: 통계청 [출판자와 배포자].
- Abikoff, H., Alvir, J., Hong, G., et al, (1987). Logical Memory subtest of the Wechsler Memory Scale: Age and education norms and alternate-form reliability of two scoring systems. *Journal of Clinical and Experimental Psychology*, 11, 783.
- Bleecker, M. L., Bolla-Wilson, K., Agnew, J., et al. (1988). Age-related sex differences in verbal memory. *Journal of Clinical Psychology*, 44, 403-411.
- Bolla-Wilson, K. & Bleecker, M. (1986). Influence of verbal intelligence sex, age, and education on the Rey Auditory-Verbal Learning Test, *Developmental Neuropsychology*, 2, 203-212.
- Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a new verbal memory test with six equivalent forms. *The Clinical Neuropsychologist*, 5, 125-142.
- Christensen, H., Hardzi-Parlovic, D., & Jacomb, P. (1997). The psychometric differentiation of dementia from normal aging: A meta-analysis.

- Psychological Assessment*, 3, 147-155.
- Craik, F. I. M. (1990). Changes in memory with normal aging: A functional view. In R. J. Wurtman, et al. (Eds.), *Advances in neurology* (Vol. 51): *Alzheimer's disease*. New York: Raven Press.
- Crossen, J. R. & Weins, A. N. (1994). Comparison of the Auditory-Verbal Learning Test(AVLT) and California Verbal Learning Test(CVLT) in a sample of normal subjects. *Journal of Clinical and Experimental Neuropsychology*, 16, 75-90.
- Cummings, J. L. & Benson, D. F. (1992). *Dementia: A Clinical approach*. Boston: Butterworth's.
- Delis, D. C., Butters, N., Cermak, L. S., Kramer, J. J., Massman, P. J., & Salmon, D. P. (1991). Profiles of demented and amnesic patients on the California Verbal Learning Test: Implications for the assessment of memory disorders, *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 3, 19-26.
- Dikmen, S., McLean, A., Mechamer, J., & Temkin, N. (1990). Neuropsychological recovery in patients with moderate to severe head injury: Two-year follow-up. *Journal of Clinical and Experimental Neuropsychology*, 12, 507-519.
- Eslinger, P. J., Benton, A. L., & Damasia, A. R. (1984). *The Iowa Screening Battery for Mental Decline*. Iowa City, IA: University of Iowa.
- Ferraro, A., & Jervis G. A. (1941). Alzheimer's disease: An attempt at establishing the adult type of the disease. *Psychiatric Quarterly*, 15, 3-16.
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101, 171-191.
- Howell, D. C. (1987). *Statistical methods for psychology*. 2nd Ed. Boston: PWS-KENT.
- Hulicka, I. M. (1966). Age differences in Wechsler Memory Scale scores. *Journal of Genetic Psychology*, 109, 135-145.
- Iverson, D. J. (1977). The Wechsler Memory Scale; Preliminary findings toward an Australian standardisation. *Australian Psychologist*, 12, 303-312.
- Iverson, D. (1986). Anna Thompson and the American Liner New York: Some normative data. *Journal of Clinical and Experimental Neuropsychology*, 8, 317-320.
- Jones, R. D., Benton, A., Paulsen, J., & Tranel, D. (1992). Differentiating dementia from 'pseudo-dementia' early in the clinical course: Utility of neuropsychological tests. *Neuropsychology*, 6, 13-21.
- Kasznik, A. W. (1986). The neuropsychological of dementia. In I. Grant & K. M. Adams(Eds). *Neuropsychological assessment of neuropsychiatric disorders*(pp. 172-220). New York. Oxford University Press.
- Kasznik, A. W., Poon, L. W., & Riege, W. (1986). Assessing memory deficits: An information processing approach. In L. W. Poon (Ed.), *Handbook for clinical memory assessment of older adults*. Washington, D.C.: American Psychological Association.
- Knopman, D. S., & Ryberg, S. A. (1989). A verbal memory test with high predictive accuracy for dementia of the Alzheimer's type. *Annals of Neurology*, 46, 141-145.
- Kramer, J. H., Brandt, J., Delis, D. C., & Levin, B. E. (1988). Differentiation of Alzheimer's,

- Huntington's and Parkinson's Disease patients on the basis of verbal learning characteristics. *Neuropsychology*, 3, 111-120.
- Lezak, M. D. (1995). *Neuropsychological assessment*. (3th ed). New York: Oxford University Press.
- Loring, D. W., Lee, G. P., & Meador, K. J. (1989). Issues in memory assessment of the elderly. In F. J. Pirozzolo(Eds.), *Clinics in Geriatric Medicine* (Vol 5., No. 3). Philadelphia: W. B. Saunders.
- Margolis, R. B., & Scialfa, C. T. (1984). Age differences in Wechsler Memory Scale performance. *Journal of Clinical Psychology*, 40, 1442-1449.
- McCarty, S. M., Logue, P. E. & Siegler, I. C. (1982). Cross-sectional and longitudinal patterns of three Wechsler Memory Scale subtests. *Journal of Gerontology*, 37, 169-175.
- Molsa, P. K., Marttila, R. J., & Rinne, U. K. (1982). Epidemiology of dementia in a Finnish population. *Acta Neurologica Scandinavica*, 65, 541-552.
- Morris, J. C., Ball, M. J., Berg, L., Grant, E. A., Mekeel, D. W., Price J. L., Rubin E. H. & Syorandt, J. M. (1991), Very mild Alzheimer's disease: Informant-based clinical, psychometric, and pathologic distinction from normal aging. *Neurology* 41: 469-478.
- O'Donnell, B. F., Drachman, D. A., Lew, R. A., & Swearer, J. M. (1988), Measuring dementia: assessment of multiple deficit domains. *Journal of Clinical Psychology*, Vol 44, No. 6. 914-923
- Rapaport, D., Gill, M. M., & Schafer, R. (1968). *Diagnostic psychological testing* (rev. ed.), Robert R. Holt (Ed.). New York: International University Press.
- Spren, O., & Strauss, E. (1997). *A compendium of neuropsychological tests*. New York: Oxford University Press.
- Storandt, M., Berg, L., Borwinick, J., Danziger W. L., & Hughes, C. P. (1984): Psychometric differentiation of mild senile dementia of the Alzheimer type. *Archives of Neurology*, 41: 497-499.
- SPSS v. 11 [Computer software] . (2001). Chicago: SPSS Inc.
- Tremmel, P. V. (2003). Women most effective leaders for today's world. *Psychological Bulletin*, Vol. 129, No. 3.
- U. S. Census Bureau--Statistical Brief: 1997--Sixty-five plus in the United States [Electronic data] . (1997). Washington, DC: U.S. Census Bureau [Producer and Distributor] .
- Verhoff, A. E., Albert, M. L., Kaplan, E., et al. (1979). *Aging and dementia in the Framingham Heart Study population: Preliminary prevalence data and qualitative analysis of visual reproductions*. Paper presented at the 7th annual meeting of the International neuropsychological Society, New York.
- Webster, J. S., Godlewski, M.C., Hanley, G. L., & Sowa, M. V. (1992). A scoring method for Logical Memory that is sensitive to right-hemisphere dysfunction. *Journal of Clinical and Experimental Neuropsychology*, 14, 222-238.
- Wechsler, D. (1945). A standardized memory scale for clinical use. *Journal of Psychology*, 19, 87-95.
- Wechsler, D. (1987). *Wechsler Memory Scale-Revised manual*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1997). *Wechsler Memory Scale- III manual*.

- San Antonio, TX: The Psychological Corporation.
- Welsh, K., Butters, N., Heyman, A., Hughes, J. P., & Mohs, R. C. (1992). Detection and staging of dementia of Alzheimer's disease: Use of the neuropsychological measures developed for the Consortium to Establish a Registry for Alzheimer's disease. *Archives of Neurology*, 49, 448-452.
- Williams, J. M. (1991). *Memory Assessment Scales. Professional Manual*. Odessa, FL: Psychological Assessment Resources.
- Zec, R. F. (1993). Neuropsychological functioning in Alzheimer's disease. In Parks, R. W., Zec, R. F., & Wilson, R. S(Eds.). *Neuropsychology of Alzheimer's disease and dementia*. New York: Oxford University Press.

원고접수일 : 2002. 11. 12

게재결정일 : 2004. 1. 29

A Standardization Study of the Story Recall Test in the Elderly Korean Population

Hyo-Jung An

Department of Psychiatry
Kangnam Mary's Hospital

Jeanyung Chey

Department of Psychology
Seoul National University

This study purported to develop a story recall test (SRT) for the elderly Korean population. It examined the reliability and validity of the test and developed norms for the elderly population. Effects of demographic variables on test performance were examined. Normative sample consisted of two hundred seventy-three elderly volunteers whose ages were between 55 and 89. Volunteers were residents of Seoul, Kyungki-do, Yesan, Daegu, Masan, and Dangjin. Education and gender proved to be significant factors influencing all SRT performances, while age was significant in some performances. Norms were developed considering significant demographic factors. The norms for the Immediate Recall were established for two education level (0-9 years, over 10 years) and gender, yielding four normative groups. The norms for the Delayed Recall and the Recognition were established for two education groups (less or equal to 9 years, and more than 10 years), two age groups (55-74 years and 75-89 years) and two gender group (male and female). In addition, memory characteristics of the normal elderly Korean population were described. The normative data of the Story Recall Test are expected to be utilized in detecting memory, learning and cognitive impairments in the elderly population. More efficient diagnostic procedures will be available for the patients with early Alzheimer's disease and other neurological conditions frequently found in the elderly population.

Keywords : Dementia, Memory, Story Recall Test, Neuropsychology, Alzheimer's disease.