

주기적 결과 모니터링을 위한 성과 측정 척도 개발 및 타당화*

연규진¹⁾ 장현진²⁾ 진아영³⁾ 송영훈⁴⁾ 조화진⁵⁾ 김민정⁶⁾ 김종민⁷⁾ 이문수⁸⁾

¹⁾서강대학교 심리학과 / 교수 ²⁾서강대학교 심리학과 / 박사수료

³⁾서강대학교 심리학과 / 석사졸업 ⁴⁾서강대학교 심리학과 / 석사수료

⁵⁾한남대학교 상담심리학과 / 부교수 ⁶⁾아주대학교 교육대학원 / 조교수

⁷⁾경인교육대학교 교육학과 / 교수 ⁸⁾경상국립대학교 교육학과 / 부교수

본 연구는 주기적 결과 모니터링(Routine Outcome Monitoring [ROM])을 위해 증상 감소와 다차원적 안녕감의 관점은 통합적으로 고려한 상담 성과 측정 척도(Outcome Measurement Scale [OMS])를 개발하고 그 타당성을 검증하고자 하였다. 주기적 결과 모니터링 활용도가 높은 기존의 성과 평정 척도(Outcome Rating Scale [ORS])의 문제점과 관련 선행연구를 검토하여 심리적 불편감, 자기개념, 친밀한 관계, 일과 역할, 그리고 전반적 만족을 측정하는 총 5문항으로 구성된 간편 성과 척도를 개발하였다. 대학생 400명을 대상으로 한 예비조사에서 문항 검토를 실시하고 탐색적 요인분석을 통해 1요인이 도출되었다. 척도의 일반화 가능성을 고려하여 전 연령대를 포함한 일반인 1,000명을 대상으로 본조사를 실시하였다. 본조사의 확인적 요인분석에서 단일 차원 모형은 양호한 모형 적합도를 보였고, 우수한 내적합치도($a = .93$)를 보였다. 본 연구에서 개발된 척도는 주기적 결과 모니터링에서 자주 활용되는 상담 성과(CORE-OM), 우울(PHQ-9), 불안(K-GAD-7)과 The World Health Organization-Five Well-Being Index(WHO-5) 척도와 유의한 상관관계를 보여 공인타당도가 확인되었다. 마지막으로 본 연구의 의의와 제한점을 논의하였다.

주요어 : 주기적 결과 모니터링, 측정기반치료, 상담 성과, 척도 개발

* 이 연구는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 3단계 산학연협력 선도대학 육성사업(LINC3.0)의 연구 결과임.

† 교신저자 : 연규진, 서강대학교 심리학과, (04107) 서울특별시 마포구 백범로 35

Tel: 02-705-8327, E-mail: kjyon@sogang.ac.kr

 Copyright ©2025, The Korean Counseling Psychological Association
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

전 세계를 힘들게 만들었던 COVID-19 팬데믹 이후로 정신건강 문제에 대한 대중들의 관심이 높아졌다. 자연스럽게 상담에 관한 관심이 증가하여 어느 때보다 상담 전문성과 효과성에 대한 요구가 커진 상황이다. 특히 한국의 경우 상담심리의 역사가 해외에 비해 상대적으로 짧고, 전문성이 인정되는 공인된 자격증이 부족한 데다가 상담 효과성 검증이나 증진을 위한 연구가 상대적으로 많지 않다 보니, 일반인들은 상담에서 어떤 변화와 효과가 있는지를 이해하고 신뢰하기가 어려울 수 있다. 상담을 이용하는 내담자뿐만 아니라 정신건강 분야의 각종 이해 관계자들과 기관(예, 보건복지부, 학교, 병원, 기업, 군, EAP 등)에서는 상담의 성과 및 효과성 입증에 대한 요구가 커지고 있다. 더욱이 정부는 2024년부터 많은 예산을 투자하여 ‘전국민 마음투자 지원사업’을 실시하고 있고, 상담 전문자격 면제화에 대한 노력이 지속되고 있어 어느 때보다 심리 상담의 전문성을 과학적으로 입증하고 향상하기 위한 노력이 필요한 시점이다. 이러한 필요성에 따라 본 연구는 상담 과정에서 내담자의 변화를 주기적으로 모니터링하여 상담의 효과를 높이고 이를 검증할 수 있는 상담 성과 측정 도구를 개발하고자 하였다.

심리학 분야에서 상담 성과 측정은 오랜 역사가 있다(Codman, 2013). 상담자들은 상담이 내담자의 복지에 긍정적으로 영향을 미칠 수 있는 심리·행동적 변화를 불러올 수 있다고 가정해 왔으나, 상담의 성과가 쉽게 관찰되거나 항상 분명하게 나타나는 것은 아니다(De Jong et al., 2023). 특별히, Eysenck(1952)가 상담이 효과적이지 않다고 주장하여 많은 논쟁을 불러일으킨 이후, 많은 연구자와 실무자들이 상담 성과나 효과에 대한 측정과 이해에 관심

을 보이기 시작했으며, 이후 Smith와 Glass(1977)의 메타분석 연구를 포함한 다수의 연구(예, Wampold, 2013)를 통해 상담의 효과가 있음이 밝혀졌다. 상담 성과 연구에 대한 역사적이고 종합적인 문헌 고찰은 본 연구의 범위를 벗어나지만, 최근 연구 경향을 살펴보면 실무 현장에서 성과를 측정하고 상담 과정 모니터링에 이를 활용하는 방향으로 나아가고 있다(Castonguay et al., 2013; Lambert, 2013에서 재인용). 이러한 발전은 상담이 실제 현장에서 효과적으로 적용되고, 그 효과성이 입증되어야 한다는(김선경, 왕은자, 2016) 현장에 기반한 증거(Practice-Based Evidence [PBE]) 접근에서 비롯된 것이며, 이는 상담심리학의 과학자-실무자 모델에서부터 강조해 왔던 과학과 현장을 연결하려는 시도로 이해할 수 있다(Barkham et al., 2010).

상담 성과는 내담자가 상담을 통해 경험하는 다양한 변화를 포괄하는 개념으로, 다양한 방식으로 정의하고 측정할 수 있다. 김계현(2000)은 상담 성과를 상담 회기 중에서 경험할 수 있는 즉각적인 변화(예, 체험 정도, 통찰, 정서 표현 등)뿐만 아니라, 회기와 회기 사이 혹은 상담을 마친 후에 나타나는 변화(예, 신념, 행동 및 정서 변화 등)를 모두 포함한다고 하였다. 최근 고흥월(2021)은 상담 성과에 대한 종합적인 이해를 위해 국내 연구를 중심으로 관련 문헌을 고찰하였고, 상담 성과를 다음과 같이 살펴볼 수 있다고 하였다. 첫째, 내담자가 지각하는 주관적 변화는 상담 성과의 중요한 요소이며, 내담자가 느끼는 편안함, 긍정적인 감정 변화, 문제의 호전 등이 이에 해당한다(김선경, 왕은자, 2016; Berman, 2020). 둘째, 내담자 일상생활에서의 어려움 해소 및 사회적 기능 향상이다. 상담 개입으

로 내담자의 사고, 행동, 및 정서가 변화하면서 결과적으로 내담자의 일상생활 및 사회적 관계가 개선되면 이를 상담 성과로 간주할 수 있다(김계현, 2000). 셋째, 연구나 상담 현장에서 사용하는 객관적 심리평가 도구(예, MMPI, BDI 등)를 활용했을 때, 상담 전후의 지표 변화를 통해 상담 성과를 확인할 수 있다. 넷째, 내담자의 생리적 변화로 상담 후 생리적 지표(예, 자율신경계, 내분비계, 뇌 기능 등)에서 나타나는 변화를 측정하여 성과를 확인할 수 있다(이화실, 김계현, 2019). 마지막으로 고흥월(2021)은 내담자의 외현적 행동과 내면적 인지 변화(예, 인지적 판단 및 의사결정, 대처 행동 및 정서 조절의 변화 등)를 상담 성과로 간주할 수 있다고 하였다. 예를 들어, 역기능적 도식의 변화는 인지적 기능과 관련된 행동 변화 측면에서도 안정적일 수 있으며(Kellogg & Young, 2008, Lebow, 2012에서 재인용), 자존감과 자기효능감 측면(Watson et al., 2002)에서의 변화도 발생한다고 하였다.

앞서 살펴보았듯이 상담 성과는 다양한 관점에서 고려할 수 있는데, 2000년대 초부터 심리검사를 활용하여 내담자의 상담 성과를 주기적으로 평가하고 이를 상담자에게 피드백하는 주기적 결과 모니터링(Routine Outcome Monitoring; [ROM])의 유용성이 전 세계적으로 강조되고 있다(Overington & Ionita, 2012). ROM은 세 가지 핵심 요소로 구성되는데 첫째, 상담 과정 전반에 걸쳐 내담자의 상태(예, 증상, 기능 등)와 관련해서 내담자가 보고한 결과를 주기적으로 수집하고, 둘째, 내담자가 보고한 상담 성과 및 시간에 따른 변화를 공유하고, 셋째, 내담자의 데이터와 상담자의 임상적 판단을 바탕으로 내담자와 상담 목표를 함께 협력하여 결정하는 과정(Lewis et al., 2019; Oslin

et al., 2019)이 포함된다. ROM은 이러한 세 가지 핵심 요소를 상담 과정에 통합하여 치료 경과에 대한 피드백을 적극 활용하여 상담의 성과를 향상하는 것을 목표로 한다. 이와 같은 접근은 ROM 이외에도 다양한 표현으로 사용되고 있는데, 예를 들어, 미국심리학회(American Psychological Association)에서는 Measurement-Based Care(Scott & Lewis, 2015)로 지칭하며, 그 외에도 Feedback Informed Treatment(예, Delgadillo et al., 2018), Routine Outcome Monitoring Feedback(예, De Jong, 2016) 등으로 표현된다.

ROM의 효과를 검증하는 다수의 연구가 진행되었는데(Lambert et al., 2001; Miller et al., 2005; Schiepek & Aichhorn, 2013), 최근 진행된 De Jong 등(2021)의 메타분석 연구에서는 58개의 연구를 검토하여 총 21,699명의 내담자를 대상으로 ROM이 증상 감소에 미치는 장단기 효과를 분석하였다. 그 결과, 상담 과정을 지속해서 모니터링하는 것은 증상의 감소, 조기 종결(drop-out) 비율의 감소 및 상담의 효과성 증진에 유의한 영향을 주는 것으로 확인되었다. 연구자들은 ROM이 특히 내담자 악화율(deterioration rate)을 감소시키고, 상담에서 좋지 않은 결과를 예방하는 데 도움이 된다고 보고하였다(Lambert & Whipple, 2018). 따라서, 미국심리학회에서는 근거기반실천(evidence-based practice)의 일환으로 ROM 활용의 중요성을 강조하고 있고(American Psychological Association, 2006), 영국의 IAPT(Improving Access to Psychological Therapies)에서는 ROM을 프로그램의 핵심 요소의 하나로 강조하고 있다.

이렇듯 해외에서는 ROM의 유용성이 입증되어 관련 연구들이 활발히 진행됐지만, 국내에는 관련 연구도 매우 제한적일 뿐만 아니라

ROM 자체가 잘 알려지지 않은 실정이다 (Hong, 2020). 국내 상담자들이 ROM을 현장에서 활용하기 위해서는 이를 학습할 기회가 필요한데, 최근 국내 대표적인 학회인 한국상담심리학회(2024년 학술대회)와 한국상담학회(2024년 추계 연수회)에서 ROM에 대한 소개와 교육이 있었으나, ROM을 체계적으로 학습할 기회가 거의 없었고, 관련 국문 서적이 전혀 없어 현장에서는 ROM에 대한 인식 자체가 부족한 상황이다. 또한, 국내에는 ROM 활용을 목적으로 개발된 척도가 제한적인데, 성과 평정 척도(Outcome Rating Scale [ORS]; Miller et al., 2003)를 변안, 타당화한 안성식 등(2020)은 연구 목적이 아닌 국내 실무자들이 ROM을 더욱 쉽게 활용할 수 있는 짧은 성과 측정 척도의 필요성을 강조하였다. 따라서 본 연구에서는 ROM에서 활용할 수 있는 짧은 상담 성과 측정 척도를 개발하고자 하며, 이는 국내 상담자들이 ROM을 시행하는 과정에서 사용할 수 있는 척도의 선택지를 넓히는 데 도움이 될 것이다.

해외에서는 상담 과정 및 성과를 모니터링하기 위해 다양한 척도를 사용하고 있으나, 다수의 척도는 내담자의 심리적 증상 변화에 초점을 두는 경향이 있다. 해외에서 개발되어 널리 쓰이는 증상 관련 척도에는 Beck Depression Inventory(BDI; Beck, 1972), Hamilton Depression Rating Scales(HDRS; Hamilton, 1960), Symptom Checklist(SCL; Derogatis et al., 1976) 등이 있다. 국내에서도 주로 이러한 증상 중심의 해외 척도들을 변안, 타당화하여 사용하고 있는데, 최근 국가 차원에서 시행하고 있는 '전국민 마음투자 지원사업'에서도 심리적 증상을 측정하는 척도들을 다수 안내서에 포함하였다(보건복지부, 2024). 구체적으로, Patient

Health Questionnaire-9(PHQ-9; Spitzer et al., 1999), Generalized Anxiety Disorder-7(GAD-7; Spitzer et al., 2006), 자살 및 자살 위험성 설문지(The P4 Screener; Dube et al., 2010), The Primary Care PTSD Screen for DSM-5(PC-PTSD-5; Prins et al., 2016) 등의 척도가 이에 해당한다. 예를 들어, PHQ-9의 경우 “기분이 가라앉거나, 우울하거나, 희망이 없다고 느꼈다”와 같은 내용으로 구성되어 있으며, 주요우울장애의 선별과 증상 감소 확인을 위해 심리적 증상에 초점을 두고 평가하게 되어있다. 영국의 IAPT에서는 성과를 확인하기 위한 주요 척도로 PHQ-9와 GAD-7를 사용했는데, IAPT의 경우 특별히 우울증과 불안장애에 초점을 둔 프로그램이었다. 이처럼 치료 대상과 치료 목표의 초점이 분명한 경우, 증상 중심의 척도는 임상군 선별 및 성과를 측정하는 데 유용하다.

하지만, ROM에서 특정 증상을 측정하는 척도(예, PHQ-9, GAD-7)를 활용하는 것은 다양한 증상과 문제를 호소하는 내담자들을 대상으로 상담을 진행할 때 어려움이 발생할 수 있다. 예를 들어, 상담자는 다양한 척도들을 숙지해야 하고 사례마다 내담자 증상에 따라 적절한 척도를 매번 선택해야 하는 부담이 생길 수 있다. De Jong 등(2023)은 ROM 시행에서 발생하는 상담자의 어려움 중의 하나로 척도 선택에 대한 불확실성을 언급한 바 있다. 또한, 다양한 척도가 사용되는 경우 성과를 판단할 수 있는 일관된 기준을 갖기 어렵기 때문에 상담 진행 과정에서 사례마다 치료 경과를 비교하여 파악하는 것이 어려워지고, 결과적으로 특정 기관이나 사업의 성과 전반을 모니터링하고 평가하는 기준을 세우기 어려워질 수 있다. 또한, 임상적 증상(예, 우울, 불안)을 경험하지 않으면서 대인관계, 자녀양육 혹

은 진로고민 등의 문제로 내방하는 경우, 상담을 통해 다양한 측면(예, 자기이해, 관계회복, 진로결정)에서 도움을 받았다 하더라도 증상 중심의 척도를 활용하여 반복적으로 측정하는 것은 상담에서의 성과를 적절히 측정하지 못하게 될 가능성이 있다. 지속적인 우울증을 경험했던 내담자들의 치료 성과에 대한 연구(Cummergen et al., 2022)에서도 내담자들은 상담 성과로 증상 회복 이상의 것들을 중요시한다고 보고했고, 증상 중심의 측정 방법들은 내담자 관점에서 표현된 회복(예, 개인적 균형, 자신감, 대인관계적 조화)을 잘 반영하지 못하게 된다고 하였다. 요약하자면, 증상 감소는 상담 성과를 측정하는 중요한 요소이지만 그 외의 측면들도 성과로 고려할 필요가 있으며, ROM에서 다양한 내담자들을 대상으로 증상 중심 척도를 사용할 때는 앞서 언급한 사례별(또한 기관 및 사업별) 상담 성과 비교의 어려움이 발생할 수 있다는 것을 인지할 필요가 있다.

이러한 증상 중심 척도의 대안으로 전반적인 안녕감과 삶의 질에 초점을 둔 척도나 다영역을 측정하는 척도를 활용하여 상담 성과를 측정할 수 있다. 안녕감과 삶의 질을 고려한 대표적인 척도에는 Quality of Life Index(QOL; Spitzer, 1981)와 WHO Psychiatric Disability Assessment Schedule(World Health Organization, 1988)과 같은 척도들이 있다. 일부 연구자들은 단순히 증상의 개선을 넘어 선 개인적, 사회적, 직업적 기능에서의 변화를 더 의미 있게 보아야 한다고 주장했다(Ogles, 2013; Pavot & Diener, 2004; World Health Organization, 1946). 안녕감을 평가하는 척도들의 경우, 내담자의 증상을 민감하게 파악하기는 어려울 수 있지만 다른 장점이 존재하는데,

예를 들어, 증상 감소에 국한하여 내담자를 평가하기보다 삶의 여러 측면에서의 긍정적인 부분을 파악하고 개입할 수 있다. 또한, 관련 연구에 의하면 안녕감이 증가하면 술이나 담배를 줄이고, 운동량이 증가하는 등 좋은 습관들을 더 많이 개발하게 되며, 이는 개인의 목표 달성과 사회적 관계 향상, 업무에서의 성과로 이어질 수 있다고 한다(Diener et al., 2018; Holt-Lunstad et al., 2010, 2015). 즉, 심리적 증상이 내담자 존재 자체가 아닌 일부라는 점을 고려하면(문현미, 2005), ROM에서 반복적으로 증상만을 파악하기보다 안녕감을 함께 고려할 경우, 내담자 삶을 거시적인 관점에서 조망하여 성과를 평가할 수 있다는 장점이 있다. 이에 본 연구에서는 증상과 관련된 심리적 불편감의 완화뿐만 아니라 내담자의 안녕감 관점은 측정하여 회복과 성장이라는 긍정적 측면을 함께 고려하는 통합적 관점에서의 상담 성과 척도를 개발하고자 하였다.

한편, 상담 성과를 측정하는 다영역 척도에는 Clinical Outcomes in Routine Evaluation-Outcome Measures(CORE-OM; Evans et al., 2000), Outcome Questionnaire-45(OQ-45; Lambert et al., 2004)와 Behavioral Health Measure-20(BHM-20; Kopta & Lowry, 2002) 등이 있다. CORE-OM은 연구자뿐만 아니라 현장에서 활동하는 상담자들 모두 사용할 수 있는 상담의 핵심 성과를 측정하고자 하는 취지로 개발되었는데(Evans et al., 2000), 국내에서는 김선경과 왕은자(2016)가 성인 내담자 600명을 대상으로 타당화하였다. CORE-OM은 전반적인 안녕감과 심리적 기능을 다차원적으로 평가하기 위한 34문항으로 구성되어 있고, 다음 네 가지의 주요 영역을 측정한다(Evans et al., 2000). 첫째, 안녕감(well-being) 영역은 개인의 전반적인 심리적 만

족도와 정서적 건강 상태를 평가한다. 둘째, 문제 증상(problems/symptoms) 영역은 우울, 불안, 신체화 증상 등 심리적 고통의 정도를 측정한다. 셋째, 기능(functions) 영역은 일상생활에서의 대인관계 및 사회적 역할 수행 능력을 평가한다. 마지막으로, 위험(risks) 영역은 자해 및 타해와 관련된 위험 행동을 측정한다(Evans et al., 2000). Lambert 등(2004)이 개발한 OQ-45의 경우, 다음과 같은 세 가지 주요 영역을 측정하고 45문항으로 구성되어 있다. 첫째, 증상 고통(symptom distress) 영역은 주로 우울과 불안을 평가한다. 둘째, 대인관계 문제 (interpersonal relations) 영역은 가족 관계, 결혼, 사회적 고립, 갈등 등을 포함하여 내담자의 대인관계 전반을 평가한다. 셋째, 사회적 역할 수행(social role performance) 영역은 직업, 학업, 일상생활에서의 적응과 역할 수행 능력을 측정한다(Lambert et al., 2004). 본 연구에서 이러한 다영역 척도들의 하위 요인을 참고하여 증상 완화뿐만 아니라 관계 및 사회적 기능 측면을 고려한 상담 성과 측정 척도를 개발하고자 하였다.

현장에서의 실제 활용도를 높이고 주기적인 모니터링이 가능하게 하려면, 긴 문항의 검사보다 짧은 문항의 간편한 검사를 개발하는 것이 필요하다. 앞서 설명했던 다영역 검사의

경우 다양한 측면을 측정할 수 있다는 장점이 있지만, 검사의 문항 수가 많기 때문에 매회 기 혹은 자주 시행하기에는 현실적인 어려움이 있다(De Jong et al., 2023). 현재 해외에서 적은 문항 수로 개발되어 ROM으로 활용되고 있는 척도에는 ORS가 있다. 이는 미국의 약물남용 · 정신건강서비스국(Substance Abuse & Mental Health Service Administration [SAMHSA])에서 근거기반실천으로 인정한 PCOMS(Partners for Change Outcome Management System)에서 사용하고 있는 척도이다. ORS의 경우 4문항으로 구성되어 있고, 수평선 상에서 해당 정도를 표시하는 직관적인 방식을 취하고 있어 매우 짧은 시간 내 응답이 가능하다. Miller 등(2003)에 의하면, 상대적으로 문항 수가 많은 OQ(Lambert et al., 2004)에 비해 ORS의 경우 상담에서의 지속적인 사용률이 더 높은 것으로 보고되었고, 구체적으로 OQ는 실시 1년 후에 25%가 사용했지만, ORS의 경우 89%가 사용했다고 보고하였다. 또한 ORS는 증상에 초점을 두기보다 개인의 전반적 기능과 안녕을 측정함으로써 증상 감소 이외의 내담자 관점에서 표현된 회복을 모니터링할 수 있다는 장점도 갖추고 있다. 이러한 장점으로 인해, ORS는 국내에서 안성식 등(2020)에 의해 타당화되었고, 이화실과 이해은(2016)의 연구에서

표 1. 안성식 등(2020) 및 이화실과 이해은(2016) 연구에서 번안된 ORS 문항

문항	안성식, 심은정, 박승민(2020)	이화실, 이해은(2016)
1	개인적으로 (개인적인 안녕감)	개인적으로 어떤 시간을 보냈나요?
2	가까운 대인관계에서 (가족, 친한 사람과의 관계)	당신의 대인관계는 어땠어요?
3	사회생활에서 (직장, 학교, 교우관계)	직장, 학교, 교회 등 집 밖에서의 당신의 삶은 어땠어요?
4	전반적으로 (전반적인 안녕감)	당신의 전반적 삶에 대한 점수는 몇 점인가요?

변안되어 사용된 바 있다.

하지만, ORS는 국내 현장에서 도입해서 활용하기에는 몇 가지 제한점이 있다. 첫째, ORS는 해외에서 개발된 짧고 모호한 문항 특성 때문에 수렴자가 원문의 의미를 정확히 이해하고 응답하기 어려운 점이 있다. 앞서 언급했던 국내 두 연구에서는 ORS의 문항을 표1과 같이 다르게 변안하였는데, 사용하는 상담자와 내담자마다 각자 다른 의미로 해석할 여지가 있다. 예를 들어, 원척도의 1번 문항인 'Individually(personal well-being)'는 안성식 등(2020)의 연구에서는 '개인적으로(개인적인 안녕감)', 이화실과 이해은(2016)의 연구에서는 '개인적으로 어떤 시간을 보냈나요?'로 변안하여 사용자에 따라 다른 의미로 해석될 수 있다. 2번 문항인 'Interpersonally' 역시 각각 '가까운 대인관계에서'와 '당신의 대인관계는 어땠나요?'라고 달리 표현하고 있는데, 전자의 경우 친밀한 관계를 후자의 경우 일반적인 관계를 묻는 것으로 해석될 가능성이 있다. 3번 문항 역시 안성식 등(2020)의 경우 교우관계를 포함한 사회생활을 묻지만, 이화실과 이해은(2016)의 경우 집 밖에서의 삶에 관해 묻고 있어 같은 내용을 측정한다고 보기는 어려울 수 있다. 또한, 2번 문항과 3번 문항의 경우 대인 관계와 사회적 관계의 의미가 중첩되는 부분이 있어 평가 시 내담자들이 혼란스럽게 인식하는 경우가 있으며, 실제 안성식 등(2020)의 연구에서 통계 결과를 바탕으로 두 문항이 관계를 중심으로 공통적인 부분을 공유하고 있음을 한계로 언급하였다. 둘째, ORS는 내담자들이 우울, 불안, 스트레스 같은 증상으로 불편감을 경험하고 있을 때 이를 직접적으로 측정할 수 있는 문항이 없어 증상 완화 및 변화를 적절히 포착하지 못할 가능성이 있다. 셋째,

ORS는 개인 상담자가 아닌 다수가 함께 운영하는 상담센터, 기업 혹은 기관에서 사용하게 되는 경우 사용 비용을 지불해야 하는 저작권이 있는 척도이기 때문에(Miller, n.d.) 큰 기관이나 대규모 사업에서 지속해서 활용하기 어려울 수 있다. 넷째, ORS는 시각적 아날로그 척도(visual analogue scale)로, 10 cm의 수평선 위에 응답자가 한 지점을 표시하면 그 길이(cm)를 자로 쟤어 점수화하는 방식을 취하고 있다(Miller & Duncan, 2004). 하지만, 이 방식은 리커트 척도에 비해 실제 상담 과정에서 채점하고 점수화하는 것이 번거롭고 시간이 더 소요되는 단점이 있다.

이에 본 연구는 기존의 ORS 척도의 단점을 보완하고 앞선 논의를 바탕으로 국내 상담 현장에서 주기적인 상담 성과 모니터링을 위해 보다 유용하게 활용할 수 있는 상담 성과 측정 척도(Outcome Measurement Scale [OMS])를 개발하고자 하였다. 측정 영역으로는 기존의 상담 성과에서 중요시했던 증상 감소의 관점뿐만 아니라 다영역 척도에서 공통으로 포함하고 있는 관계 및 사회적 기능 측면을 고려하여 다차원적 안녕감의 관점을 포괄하고자 하였다.

본 연구에서 개발하고자 하는 OMS에는 앞서 살펴본 기존 척도들의 한계들을 개선하고자 다음의 내용을 반영하고자 한다. 첫째, 내담자의 심리정서적 어려움(예, 우울, 불안, 스트레스) 정도를 측정함으로써 상담을 통해 내담자의 심리적 불편감의 변화를 확인할 수 있도록 하고자 한다. 이는 증상 완화를 상담 성과로 바라보는 기존의 관점을 고려한 것으로 특정 증상을 호소하는 내담자들이 해당 문항을 통해 자신의 증상과 관련된 변화를 표시하고, 증상 감소 측면에서의 상담 성과를 포착

할 수 있도록 하기 위함이다.

둘째, OMS는 안녕감에 대한 최근 이론적 논의를 반영하여 안녕감의 다차원적 특성과 기능적 효용성을 고려하고자 한다. Diener (1984)가 개인의 행복에 관한 과학적 탐구의 필요성을 강조하며 삶의 만족, 궁정과 부정 정서의 균형 등을 강조하는 주관적 안녕감 (subjective well-being) 개념을 제안한 이후, 지금 까지 다양한 측면에서 안녕감에 대한 논의가 활발히 이루어져 왔다. 먼저 Ryff(1989)는 궁정적 기능, 자기실현으로 성취되는 선함을 강조하는 자기실현적(eudaimonic) 안녕감, 즉, 심리적 안녕감(psychological well-being)을 제안하였다. 심리적 안녕감은 주관적 안녕감의 쾌락적 특성과 달리 개인의 자기실현, 성장 및 자기 수용 등을 강조한다. 이후 Keyes(2006)는 심리적 안녕감에 사회적 측면이 빠져 있다고 비판하며, 개인의 건강에서 매우 중요한 인간관계, 그리고 사회적 기능의 질을 측정하는 사회적 안녕감(social well-being) 개념을 제안하였다. 심리적 안녕감이 개인 사적인 삶의 관점에 초점을 두었다면, 사회적 안녕감은 사회구조 혹은 공동체에 속했을 때 당면하게 되는 관계, 역할, 일 등의 사회적, 기능적 측면에 초점을 두는 개념이라 할 수 있다. 따라서, 개발하고자 하는 OMS에서는 안녕감과 관련한 앞선 이론적 논의를 고려하여 자기 인식 및 수용 등의 개인 성장을 강조하는 심리적 안녕감, 사회적 관계의 질 및 일과 역할을 고려한 사회적 안녕감, 삶의 만족과 정서적 만족을 강조하는 전반적인 주관적 안녕감의 세 측면을 모두 반영하고자 한다. 이러한 시도는 상담 성과 요인을 증상 감소에 한정을 짓기보다는 다양한 성과 요인, 특히 상담에서의 변화를 유지할 수 있는 개인 내적 성장이 포함되어야 함을

강조하는 최근의 주장(예, 이성욱, 김수임, 2024)과 개인적, 사회적, 직업적 기능에서의 변화를 중요시해야 한다는 주장(Ogles, 2013; Pavot & Diener, 2004; World Health Organization, 1946)에 근거한 것이다.

셋째, 한국 내담자들이 쉽게 이해할 수 있는 문항을 개발하고자 한다. 이를 위해 우선, 자기(self) 및 관계에 대한 인식이 서구문화와는 상이한 우리 문화(Triandis, 1994)의 특징이 반영될 수 있는 언어를 사용하여 문항을 개발하고자 하였다. 전술한 바와 같이 연구자나 상담자가 아닌 내담자들의 관점에서 이해하기 쉬운 표현을 사용하고, 자기실현적 측면의 안녕감이나 사회적 안녕감을 측정하기 위해서 번역된 학문적 용어를 사용하기보다는 한국 맥락에서 내담자가 ‘자기’와 ‘친밀한 사회적 관계’, ‘역할’을 떠올릴 수 있는 일상생활에서 쉽게 접할 수 있는 표현으로 문항을 구성하고자 한다.

정리하면, 본 연구는 국내 상담 장면에서 내담자의 변화와 상담 성과를 주기적으로 측정하기 위해 증상감소와 안녕감이라는 두 측면을 모두 반영한 통합적 관점에서의 상담 성과 측정 척도(OMS)를 개발하고자 한다. ROM을 위한 국내 척도가 부족한 상황에서 개발된 척도를 통해 ROM을 시행하고 객관적 데이터를 수집하여, 내담자들의 변화와 상담의 성과를 가시적으로 파악함으로써 상담의 질을 높이는 데 기여하고자 한다. 구체적인 연구 문제는 다음과 같다. 첫째, 개발한 OMS의 요인 구조는 어떠한가? 둘째, OMS의 신뢰도와 타당도는 어떠한가? 본 연구에서는 OMS의 심리적 불편감, 자기개념, 친밀한 관계, 일과 역할, 전반적 만족의 5개 영역을 상담 성과의 단일 요인 구조로 가정하였고, 안녕감을 측정하는

WHO-5 index와는 정적 상관을, 부정적 정서를 측정하는 우울 PHQ-9, 불안 GAD-7과는 부적 상관을, 그리고 다형역으로 구성된 상담 성과 측정 척도 CORE-OM과도 부적 상관을 보일 것을 가정하였다.

방 법

문헌조사 및 선행연구 분석

상담의 성과 측면에서 내담자 변화와 성과를 모니터링할 수 있는 상담 성과 측정 척도를 개발하기 위해 8명의 연구진(상담심리학 교수 3인, 방법론 교수 2인, 박사 과정 1인, 석사 졸업 1인, 석사 과정 1인)이 참여하였다. 연구진들은 ROM과 관련한 문헌을 고찰하고 논의하였으며, 나아가 기준 ROM에서 활용하고 있는 다수의 척도를 검토하고, 증상 감소와 다차원 안녕감을 통합적으로 반영한 간단한 성과 척도를 개발하고자 하였다.

문항의 개발을 위해 선행연구 분석 과정에서 ROM에서 활용되고 있는 다형역 척도뿐만 아니라, 짧은 문항으로 구성된 안녕감을 측정하는 척도들(예, WHO-5, ORS)을 면밀히 검토하였다. 특별히, 미국심리학회에서 근거기반실천으로 인정한 PCOMS에서 사용하고 있는 ORS척도의 국내 번안된 버전(예, 안성식 등, 2020; 이화실, 이혜은, 2016)의 문제점을 검토하여 문항을 개발하였다.

포커스 그룹 인터뷰

개발된 문항의 적절성을 확인 및 보완하기 위해 연구팀에서 3차례의 포커스 그룹 인터뷰

를 진행하였다. 상담 및 임상 전공 대학원생 12명(1급 상담전문가 자격증 소지자 3명 포함)을 대상으로 1, 3차 포커스 그룹 인터뷰를 진행하였고, 학부 대학생 8명을 대상으로 2차 포커스 그룹 인터뷰를 진행하였다. 각 포커스 그룹 인터뷰는 약 60~70분씩 진행되었다. 포커스 그룹에서는 개발된 문항들을 읽고 문항의 표현을 검토하고, 문항의 측정 의도가 어떻게 이해되는지, 응답자 관점에서 각 문항에 대해 어떻게 응답하게 될 것인지에 대해 질문하였다. 포커스 그룹에서 사용했던 질문 문항의 예시를 다음에 소개하고, 사용했던 포커스 그룹 차수를 팔호에 표시하였다. “문항을 읽었을 때 직관적으로 이해가 되는지, 혹은 의미가 모호한 부분이 있는지 설명해 주세요”(1차, 2차, 3차), “특정 단어나 표현이 이해하기 어렵거나 해석을 달리할 가능성이 있는 부분이 있는지 말씀해 주세요”(1차, 2차, 3차), “상담자 입장에서 생각했을 때, 각 문항을 통해 무엇을 측정하고자 하는지 그 의도를 짐작해서 설명해 주세요”(1차, 3차), “각 문항이 상담을 통한 성과를 측정하는데 적절하다고 생각하는지 말해 주세요”(1차, 2차), “자신이 상담을 받는다고 가정했을 때, 각 문항을 읽고 어떤 점들을 고려해서 응답할 것 같은지 설명해 주세요”(1, 2차), “[각 문항에 대한 연구자의 의도를 설명한 후] 연구자가 의도한 바와 다르게 해석될 가능성이 있는 문항이 있다면, 어떻게 수정하면 좋을지 제안해 주세요”(3차). 이러한 질문을 바탕으로 포커스 그룹을 진행하여 문항들을 수정 보완하였다. 예를 들어, 1번 문항의 초기 예비 문항에서는 ‘심리적 불편감’이라는 표현을 사용하였을 때, 학부생들의 경우 이 표현의 의미가 잘 다가오지 않는다고 하여 ‘심리-정서적 어려움’ 표현을 사용하게 되었다.

4번 문항의 경우, ‘역할’이라는 단어를 포함하여 한국 문화적 맥락에서 과업이나 직장생활과 관련된 측면을 잘 파악할 수 있도록 문항을 만들었다. 또한, 2번 “자신 스스로를 바라봤을 때 어떠셨나요?” 문항의 경우, 참가자들이 자기인식, 자아존중감, 자기수용 등의 관점에서 응답하게 된다는 것을 파악하여 연구자의 의도가 잘 반영되고 있음을 확인하였다. 또한, 2번과 5번 문항의 경우, 유사한 측면이 있으나 2번의 자기개념 측면의 문제와 5번의 삶 전반의 만족도가 구별되는 상담 사례들이 종종 있다는 의견을 고려하여 문항에 모두 포함하였다. 이러한 과정을 통해 문구를 검토, 수정 및 합의하는 절차를 거쳤고, 이를 통해 내담자들에게 의미가 분명히 전달되도록 문항을 개발하였다.

문항개발

OMS는 지난 1주간의 경험을 탐색하며, 내담자들이 경험할 수 있는 증상과 관련된 심리적 불편감을 고려하기 위해 1번 문항 “심리-정서적 어려움(예, 우울, 불안, 스트레스 등) 측면에서 어떠셨나요?”를 개발하였다. 이는 주관적 불편감 척도(Subjective Units of Disturbance Scale [SUDS]; Wolpe, 1990)의 성격을 반영하여 증상과 관련된 내담자의 심리적 불편감의 변화를 포착하고자 하였다. 다음으로 자기개념(self-concept)을 탐색하기 위해 2번 문항 “자신 스스로를 바라봤을 때 어떠셨나요?”를 개발하였고, 이를 통해 자기개념, 자기수용 등의 관점을 탐색할 수 있도록 하였다. Rogers(1959)의 인간중심이론에서는 자기개념을 강조하였으며, 이는 개인의 심리적 성장 및 안녕감에 중요한 역할을 하는 것으로 보았다. 자신의 조건이나

경험을 있는 그대로 받아들이는 자기수용(self-acceptance)은 자신이 통제할 수 없는 문제에 대해 과도하게 반응하지 않도록 하며(Ryff, 1989), 안정적인 정서와 심리적 건강에 긍정적 영향을 미치는 것으로도 알려져 있다(김광웅, 2007; Ellis & Robb, 1994). 즉, 자기개념과 자기수용은 개인의 심리적 안녕감을 나타내는 지표로(Ellis & Harper, 1997), 이는 치료적 효과를 평가하는 데 중요한 요인이 될 수 있다. 사회적 안녕감 관점을 고려하여, 친밀한 관계를 측정하기 위해 3번 문항 “당신과 가까운 사람들과의 관계에서 어떠셨나요? (예, 가족, 연인, 친한 친구 등)”를 개발하였고, 전반적인 대인 관계가 아닌 내담자와 친밀한 대인관계에 대해 측정하고자 하였다. 추가로 직업적 기능 측면을 측정하기 위해 4번 문항 “당신에게 주어진 일과 역할 측면에서 어떠셨나요? (예, 직장, 학교, 가정 등)”를 개발하였고, 일과 ‘역할’이라는 표현을 포함함으로써 한국의 사회문화적 맥락을 고려하였다. 집단주의 문화에서는 역할이 개인주의 문화에 비해 더욱 강조되고, 단순히 일이나 업무 자체가 아니라 공동체, 직장, 가정 등 다양한 그룹 내에서 자신의 역할을 강조하고 정체성을 정의하는 경우가 많다는 점(오현숙, 민병배, 2008; Hofstead, 2001)을 고려하여 문항을 개발하였다. 마지막으로 5번 문항 “전반적인 삶의 만족과 행복 측면에서 어떠셨나요?”를 포함하여, 앞선 문항들이 측정하지 않았지만, 내담자 관점에서 고려할 수 있는 삶 전반의 만족과 행복을 측정하여 전반적인 주관적 안녕감을 반영하고자 하였다.

응답 방식 측면에서, OMS는 상담 현장에서의 유용성과 변화에 대한 민감성을 고려하여 11점 리커트 척도(0점: 매우 나쁨, 10점: 매우 좋음)로 개발하였다. 변화에 대한 민감성은 특

별히 ROM에서 활용하는 척도에서 중요한 기준이 되는데, ROM에서는 척도를 반복적으로 활용하기 때문에 비교적 짧은 주기에 나타나는 작은 변화를 포착할 수 있어야 한다(De Jong et al., 2023). 척도의 변화에 대한 민감성은 일반적으로 척도의 응답 범위가 넓을수록 좋아지기 때문에(De Jong et al., 2023), 본 연구에서도 내담자가 경험하는 1~2주 간격의 회기 사이의 작은 변화를 포착하기 위해 11점 리커트 방식을 사용하였다.

연구 참가자 및 자료수집 절차

본 연구에서는 조사 전문 업체에 의뢰하여 예비조사와 본조사를 위해 총 1,400명의 자료를 수집하였다. 예비조사에서는 만 20~29세 대학생 400명의 자료를 수집하였으며, 본 조사에서는 전 연령대를 포함한 만 20~69세 일반 성인 1,000명의 자료를 수집하였다. 참여자들은 본 연구의 목적과 방법에 대한 설명문을 읽고 자발적으로 연구 참여에 동의한 후 설문에 참여하였다. 본 연구는 서강대학교의 생명윤리위원회(Institutional Review Board [IRB])의 승인을 받은 이후 진행하였다 (SGUIRB-A-2308-28).

예비조사에서는 성인 대학생을 대상으로 성별, 학년 및 거주지역별 인구비례 대비로 표집하였다. 참가자의 성별은 남성 200명(50.0%), 여성 200명(50.0%), 학년별로 각 100명씩 모집되었다. 현재 거주지역으로는 경기도 109명(27.25%), 서울 88명(22%), 경상남도 21명(5.25%), 부산 24명(6%), 인천 23명(5.75%), 경상북도 16명(4%), 대구 18명(4.5%), 충청남도 15명(3.75%), 전라남도 11명(2.75%), 전라북도 12명(3%), 충청북도 12명(3%), 강원도 11명

(2.75%), 대전 13명(3.25%), 광주 12명(3%), 울산 8명(2%), 제주도 5명(1.25%), 세종시 2명(0.5%)이었다. 본조사에서는 일반화 가능성을 높이기 위해, 대학생 연령대를 포함한 만 20~69세 일반 성인을 대상으로 성별, 연령 및 거주지역별 인구비례 대비로 표집하였다. 참가자의 성별은 남성 510명(51.0%), 여성 490명(49.0%)이며, 연령은 50대 237명(23.7%), 40대 213명(21.3%), 60대 207명(20.7%), 30대 174명(17.4%), 20대 169명(16.9%) 순이었다. 현재 거주지역으로는 경기도 270명(27.0%), 서울 189명(18.9%), 경상남도 62명(6.2%), 부산 62명(6.2%), 인천 60명(6.0%), 경상북도 48명(4.8%), 대구 48명(4.8%), 충청남도 39명(3.9%), 전라남도 32명(3.2%), 전라북도 31명(3.1%), 충청북도 31명(3.1%), 강원도 29명(2.9%), 대전 29명(2.9%), 광주 26명(2.6%), 울산 21명(2.1%), 제주도 13명(1.3%), 세종시 10(1.0%)명이었다.

측정 도구

상담 성과 척도 (OMS)

주기적 상담 성과 모니터링을 위해 단일요인 총 5문항의 상담 성과 척도(OMS)를 개발하였다. 본 척도는 기존의 내담자 증상에 초점을 둔 관점을 반영하여 심리적 불편감 정도를 측정하였고, 안녕감의 다차원적 관점을 반영하여 자기개념, 친밀한 관계, 사회적 기능과 역할, 전반적 만족 측면을 측정하였다. 11점 리커트 척도(0점: 매우 나쁨, 10점: 매우 좋음)로 평정하게 되어있으며, 총점이 높을수록 심리적 불편감은 적고 안녕감이 높은 것으로 해석할 수 있다. OMS의 신뢰도는 .93으로 나타나, 높은 신뢰도 기준을 충족하는 것으로 확인하였다(Nunnally, 1978; Nunnally &

Bernstein, 1994).

한국판 핵심 상담 성과 척도(The Korean Inventory of Clinical Outcomes in Routine Evaluation-Outcome Measure [K-CORE-OM])

상담 성과를 측정하기 위해 Evans 등(2000)이 개발한 핵심 상담 성과도구(Clinical Outcomes in Routine Evaluation-Outcome Measure [CORE-OM])를 김선경과 왕은자(2016)가 타당화한 한국판 핵심 상담 성과 척도를 사용하였다. 이 척도는 34문항으로 구성되어 있으며 각 문항은 상담 성과를 확인하기 위해 지난 한 주 동안 의기소침, 불안 및 스트레스, 위기, 낮은 긍정적 태도, 사회적 지지 부족에 대하여 5점 리커트 척도(0점: 전혀 그렇지 않다, 4점: 거의 항상 그렇다)로 평정한다. 합산한 총 점이 높을수록 심리적인 문제가 심각하고 낮은 기능 수준을 보일 것을 의미한다. 문항의 예로는 ‘완전히 혼자이고, 소외되었다고 느낀 적이 있다’, ‘내 문제로 인해 압도당하는 느낌이다’가 있다. 김선경과 왕은자(2016)의 연구에서 신뢰도 계수(Cronbach's α)는 .92였으며, 본 연구에서는 .95였다. 본 척도는 심리적 고통 상황에서 우울 및 불안을 일반적으로 경험하게 되는 것을 고려하여 선행연구에서 벡 우울 척도(BDI)와 벡 불안척도(BAI) 등을 사용하여 공인타당도를 확인한 바 있다(김선경, 왕은자, 2016; Barkham et al., 2001; Evans et al., 2002).

한국어판 우울증 선별도구 (PHQ-9)

우울을 측정하기 위해 Spitzer 등(1999)이 개발한 우울증 선별도구(PHQ-9)를 안제용 등(2013)이 표준화한 한국어판 우울증 선별도구 (PHQ-9)를 사용하였다. 이 척도는 9문항으로 구성되어 있으며 각 문항은 DSM-IV(Diagnostic

and Statistical Manual of Mental Disorders, Fourth Edition; American Psychiatric Association, 1994)의 주요 우울장애 진단 기준을 확인하기 위해 지난 2주 동안 얼마나 증상을 경험하였는지 4점 리커트 척도(0점: 전혀 없음, 3점: 거의 매일)로 평정한다. 합산한 총점이 높을수록 우울로 인한 문제가 심각할 것을 의미한다. 문항의 예로는 ‘매사에 흥미나 즐거움이 전혀 없다’, ‘식욕이 줄었다 (혹은) 너무 많이 먹는다’가 있다. 안제용 등(2013)의 연구에서 신뢰도 계수 (Cronbach's α)는 .95였으며 본 연구에서는 .87이었다. 선행연구에서 PHQ-9의 수렴타당도는 PHQ-2(The Patient Health Questionnaire-2 [PHQ-2]; Kroenke et al., 2003), BDI, HDRS와의 유의한 상관을 통해 확인되었다(안제용 등, 2013).

한국형 범불안장애 척도(Korean version of the Generalized Anxiety Disorder-7 [K-GAD-7])

불안 수준을 측정하기 위해 Spitzer 등(2006)이 개발한 범불안장애 척도(Generalized Anxiety Disorder-7)를 Lee 등(2022)이 타당화한 한국형 범불안장애 척도(K-GAD-7)를 사용하였다. 이 척도는 7문항으로 구성되어 있으며 각 문항은 지난 2주 동안 불안 관련 증상으로 인한 일상에서의 불편함을 알아보기 위해 4점 리커트 척도(0점: 전혀 방해 받지 않는다, 3점: 거의 매일 방해 받았다)로 평정한다. 합산한 총점이 높을수록 불안으로 인한 문제가 심각한 것을 의미한다. 문항의 예로는 ‘걱정하는 것을 멈추거나 조절할 수 없다’, ‘너무 안절부절 못해서 가만히 있기가 힘들다’가 있다. Lee 등(2022)의 연구에서 신뢰도 계수(Cronbach's α)는 .88이었으며 본 연구에서는 .93이었다. 선행연구에서 GAD-7은 불안 측정 도구인 BAI와 HAM-A

(Hamilton Anxiety Rating Scale; Hamilton, 1959), 우울 측정 도구인 PHQ-9, BDI 및 HDRS와의 상관을 통해 공인타당도가 확인되었다(Lee et al., 2022).

WHO-5 웰빙지수(World Health Organization -5 Wellbeing Index [WHO-5])

안녕감을 측정하기 위해 세계보건기구가 처음 제안한 이후 지금까지 30개 이상의 언어로 번역되어 보급되어 있는 WHO-5 웰빙 지수 척도를 사용하였다(WHO, 1998). 본 연구에서는 김현지 등(2010)이 번안하여 국내에서 활용되고 있는 문항을 사용하였다. 이 척도는 5문항으로 구성되어 있으며 각 문항은 지난 2주 간 개인이 경험한 느낌에 대하여 6점 리커트 척도(0점: 그런적 없다, 5점: 항상 그렇다)로 평정한다. 합산한 총점이 13점 미만인 경우는 우울감이 있는 상태를 의미하며 총점이 13점 이상부터는 높을수록 안녕감이 높음을 의미한다. 문항의 예로는 ‘나는 즐겁고 기분이 좋았다’, ‘나는 활발하고 활기찼다’가 있다. 기존 연구에서 신뢰도 계수(Cronbach's α)는 .83(김현지 등, 2010), .91(최성수 등, 2023)으로 보고되었으며, 본 연구에서는 .94 였다. 또한 Topp 등(2015)의 WHO-5에 대한 문헌고찰 연구에서, 본 척도의 구조적 타당도 측면에서 각 문항들이 전반적인 주관적 안녕감을 측정하는 데 기여하며, HDRS 및 HAM-A와의 유의한 상관을 통해 공인타당도를 확인하였고, DSM-IV의 우울증 선별 도구로 사용될 때의 진단 유용성 및 예측타당도가 있는 것으로 보고되었다.

분석 방법

본 연구에서는 예비조사에서의 문항 분석,

탐색적 요인분석, 본조사에서의 확인적 요인 분석과 신뢰도 분석 및 타당도 분석을 통해 척도 타당화 작업을 진행하였다. 확인적 요인 분석은 Mplus 7을 사용하였고, 그 외의 통계분석은 SPSS 23을 사용하였다. 구체적인 분석 과정은 다음과 같다. 먼저 예비조사 자료($n=400$)를 바탕으로 평균값, 분산, 문항 간 상관 등을 통한 문항 분석을 하였고, 이후 수집된 자료의 구조가 요인분석에 적합한지 확인하기 위해 KMO 표본 적합도 지수(Kaiser-Meyer-Olkin)를 확인하고 Bartlett 구형성 검정(Bartlett's test of sphericity) 결과를 살펴보았다. 적정 요인 수를 결정하기 위해 상관행렬의 고윳값과 스크리 도표를 검토하였다. 예비조사에서 수집된 자료를 바탕으로 탐색적 요인분석을 실시하였고, 본 조사에서는 일반인 대상 인구비례 대비 표집한 자료($n=1,000$)를 바탕으로 요인 구조의 적합성을 확인하기 위해 확인적 요인분석을 실시하였다. 마지막으로 신뢰도 분석과 준거타당도를 확인하기 위해 CORE-OM과 PHQ-9, K-GAD-7, 그리고 WHO-5 척도와의 상관분석을 실시하였다.

결과

문항분석과 탐색적 요인 분석

예비조사에서는 대학생 400명 자료를 바탕으로 개발한 5개의 예비 문항 분석과 탐색적 요인분석을 실시하였다. 우선 문항의 적절성 여부를 판단하기 위해 문항분석을 실시하였다. 문항분석은 크게 세 가지 과정으로 이루어졌다. 우선, 평균과 표준편차를 계산하여 극단적인 평균값을 갖는 문항이 있는지, 표준편차가

표 2. 문항 별 평균, 표준편차, 왜도, 첨도 및 상관 (n=400)

	OMS OMS	심리적 심리적 불편감	OMS 자기개념	OMS 친밀한 관계	OMS 일과 역할	OMS 전반적 만족
OMS						
OMS 심리적 불편감1		.804 ^{**}				
OMS 자기개념		.910 ^{**}	.681 ^{**}			
OMS 친밀한 관계		.803 ^{**}	.482 ^{**}	.647 ^{**}		
OMS 일과 역할		.857 ^{**}	.572 ^{**}	.726 ^{**}	.666 ^{**}	
OMS 전반적 만족		.916 ^{**}	.671 ^{**}	.849 ^{**}	.688 ^{**}	.722 ^{**}
평균						
표준편차	29.903	5.410	5.695	6.680	6.065	6.053
왜도	10.230	2.589	2.339	2.208	2.389	2.403
첨도	-.238	.010	-.346	-.451	-.380	-.388

주 1. 심리적 불편감은 “심리·정서적 어려움(예, 우울, 불안, 스트레스) 측면에서 어땠셨나요?”라는 문항으로 측정되며, 매우 나쁨(어려움): 0에서 매우 좋음(어렵지 않음): 10사이로 응답하여 점수가 높을수록 심리적 불편감이 없음을 의미함.

주 2. ^{**}p<.01

너무 작은 문항이 있는지 검토하였다(표 2). 그 결과, 평균은 5.41~6.68, 표준편차는 2.21~2.59의 범위를 보여 변별력이 떨어지는 문항은 없는 것으로 나타났다. 다음으로 각 문항의 왜도와 첨도를 살펴보았다. 모든 문항에 대해 왜도와 첨도가 각각 절대값 3과 8을 초과하지 않아 정규분포 가정이 충족되는 것으로 나타났다(Kline, 2011). 마지막으로 문항 간 상관과 문항-총점 간 상관을 살펴보았다. 문항 간 상관은 .482~.849, 문항-총점 간 상관은 .803~.916 범위로 나타났다. 일반적으로 문항 간 상관이 .30 이하 또는 .80 이상일 때, 또 문항-총점 간 상관이 .20 이하인 경우에는 양호도가 떨어진다고 판단하는데, 자기개념을 묻는 2번 문항(자신 스스로를 바라봤을 때 어

떠셨나요?)과 전반적 안녕감을 묻는 5번 문항(전반적인 삶의 만족과 행복 측면에서 어땠셨나요?)은 상관이 .849로 다소 높게 나타났다. 두 문항이 안녕감 측면에서 관련성이 있었던 것으로 추측되나, 내용적으로 전자는 자기인식 및 자기수용과 같은 자기개념을 포함한 개인적인 차원의 관점을, 후자는 삶 전반의 주관적 안녕감을 의미한다는 점에서 상담 장면에서는 상황에 따라 차별화되는 답변을 기대할 수 있을 것으로 판단하였다. 이에 개발한 5개 예비 문항을 모두 활용하여 탐색적 요인분석을 실시하였다.

탐색적 요인분석을 실시하기에 앞서 표본 자료가 요인분석에 적합한지 확인하기 위해 예비조사(n=400명) 자료를 바탕으로 KMO 값

과 Bartlett의 구형성 검정 결과를 살펴보았다. KMO 값은 .80 이상일 경우 양호하다고 보고 (Kaiser, 1974), Bartlett 구형성 검정에서는 유의 확률이 .05 미만일 때 적합하다고 본다(Bartlett, 1950). 본 연구 표본 중 대학생 400명 자료에서는 KMO 값이 .871, Bartlett 유의확률이 .001 미만으로 나타나 요인분석이 가능한 자료임을 확인하였다.

이어 적정 요인 수를 결정하기 위해 상관행

렬의 고윳값과 스크리 도표를 검토하였다. 먼저 Kaiser(1960) 규칙에 따라 상관행렬의 고윳값이 1.0을 초과하는 값을 지닌 요인 수를 살펴본 결과, 이를 만족시키는 요인은 1개로 나타났다. 스크리 검정은 스크리 도표 상에서 고윳값이 급격히 변화되어 작아지기 직전의 요인 수를 유효 요인 수로 판단하는데(Cattell, 1966), 본 연구의 표본 자료에서는 두 번째 요인부터 경사가 완만해지므로 적정 요인 수는

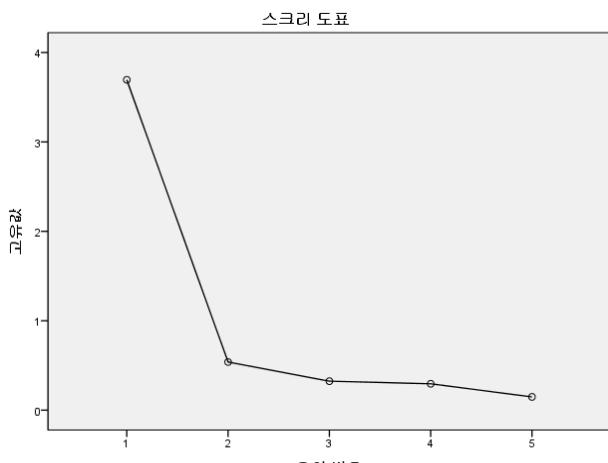


그림 1. 스크리 도표

표 3. 탐색적 요인분석 결과

문항	요인1
5. 전반적인 삶의 만족과 행복 측면에서 어떠셨나요?	.924
2. 자신 스스로를 바라봤을 때 어떠셨나요?	.914
4. 당신에게 주어진 일과 역할 측면에서 어떠셨나요? (예, 직장, 학교, 가정 등)	.798
3. 당신과 가까운 사람들과의 관계에서 어떠셨나요? (예, 가족, 연인, 친한 친구 등)	.736
1. 심리·정서적 어려움(예, 우울, 불안, 스트레스) 측면에서 어떠셨나요?	.726
고윳값	3.395
설명분산(%)	67.898

1개라고 볼 수 있다(그림 1). 이상 고윳값과 스크리 도표를 검토하여 1요인으로 고정한 뒤 탐색적 요인분석을 실시한 결과는 표 3과 같다. 각 문항의 요인부하량은 .726~.924로 나타나 일반적인 기준치인 .30을 넘었고, 1요인에 대한 설명분산은 67.898%로 나타나 적절하다고 판단하였다.

확인적 요인분석

본조사에서는 대학생 연령대를 포함한 일반인 1,000명을 인구비례 대비로 표집하였고 수집된 자료를 바탕으로 확인적 요인분석을 통해 모형 적합도를 확인하였다. 모형의 적합도를 평가하기 위해서는 χ^2 를 비롯해 여러 적합도 지수를 검토하는데, CFI가 .95 이상이면서 SRMR이 .08 이하일 경우(Hu & Bentler, 1999) 양호한 적합도를 지닌다고 보며, RMSEA가 .05 이하이면 매우 근사한 좋은 적합도를 지닌다고 본다. 앞서 탐색적 요인분석을 통해 파악

한 1요인 모형은 CFI .998, SRMR .007, RMSEA .040으로 산출되어 모형 적합도가 양호한 것으로 판단되었다(표 4).

확인적 요인분석을 통한 모수추정 결과는 표 5에 제시하였다. 모수 추정치는 모두 유의하였고, 각 문항이 1요인에 대해 .762~.916 범위의 요인부하량을 지니는 것으로 나타났다. 모든 문항이 하나의 요인에 대해 .50 이상의 부하량을 지님에 따라 수렴타당도가 확인되었다(Hair et al., 2006).

신뢰도 분석 결과

다음으로 OMS의 신뢰도 계수를 살펴본 결과는 표 6과 같다. OMS 전체에 대한 내적합치도는 .925로 나타났고, 이것은 .70 이상이면 좋다는 Nunnally(1978)의 기준을 충족하여 OMS가 신뢰로운 측정 도구임을 확인하였다. 또한 문항 삭제 시 내적합치도가 높아지는 문항은 존재하지 않았다.

표 4. 1요인 모형의 적합도

χ^2	<i>df</i>	<i>p</i>	RMSEA	CFI	TLI	SRMR
13.100	5	.0225	.040	.998	.996	.007

표 5. 확인적 요인분석을 통한 모수 추정 결과

모수	비표준화 추정치	표준오차	표준화 추정치
1요인	1 1.000		.762***
	2 1.023***	.032	.916***
	3 .897***	.032	.819***
	4 .939***	.033	.832***
	5 1.101***	.035	.902***

주. *** $p < .001$

연규진 등 / 주기적 결과 모니터링을 위한 성과 측정 척도 개발 및 타당화

표 6. 신뢰도 계수

요인	문항	문항이 삭제된		문항-척도 Cronbach's <i>a</i>
		경우	Cronbach's <i>a</i>	
	1	.924		
	2	.896		
1요인	3	.911	.925	
	4	.910		
	5	.898		

준거타당도 분석 결과

OMS의 준거타당도를 확인하기 위해 본조사 일반인 표본을 활용하여 K-CORE-OM과 PHQ-9, K-GAD-7, 그리고 WHO-5 척도와의 상관분석을 실시하였다. 먼저, OMS는 다영역 상담 성과를 측정하는 K-CORE-OM, 우울을 측정하는 PHQ-9, 불안을 측정하는 K-GAD-7과 각각 -.783, -.699, -.645을 보였고, 안녕감을 측정하는 WHO-5와는 .806의 상관을 보였다. 즉 5개 문항으로 구성된 OMS는 높은 공인타당도를 보이는 것으로 나타났다.

표 7. OMS 문항 간 및 척도 간 상관분석 결과 (*n*=1,000)

	OMS 전체	OMS 심리적 불편감	OMS 자기개념	OMS 친밀한 관계	OMS 일과 역할	OMS 전반적 만족	K- CORE- OM	PHQ -9	K-GAD -7	WHO -5
OMS 전체										
OMS 심리적 불편감		.843**								
OMS 자기개념			.916**	.699**						
OMS 친밀한 관계				.862**	.639**	.740**				
OMS 일과 역할					.867**	.626**	.771**	.696**		
OMS 전반적 만족						.912**	.698**	.830**	.740**	.739**
K-CORE-OM							-.783**	-.710**	-.728*	-.660**
PHQ-9								-.699**	-.650**	-.652**
K-GAD-7									-.645**	-.645**
WHO-5										.806**

주. ***p*<.01

논 의

본 연구는 주기적 결과 모니터링에서 활용하기 위해 증상 감소와 다차원적 안녕감의 관점과 포함하는 통합적 관점에서의 상담 성과 측정 도구를 개발하고, 그 신뢰도와 타당도를 검증하려는 목적으로 진행되었다. 이를 위해 관련 선행연구들을 검토하였으며 미국의 약물 남용·정신건강서비스국(SAMHSA)에서 근거기 반실천으로 인정한 PCOMS에서 사용하고 있는 ORS 척도와 국내 번안된 버전(예, 안성식 등, 2020; 이화실, 이해은, 2016)의 제한점을 보완하여 5문항의 성과 측정 척도(OMS)를 개발하였다.

연구의 주요 결과는 다음과 같다. 먼저, 증상 감소와 안녕감 관점을 고려한 심리적 불편감, 자기개념, 친밀한 관계, 사회적 기능과 역할, 그리고 전반적인 만족도를 묻는 5문항 11점 리커트 척도의 상담 성과 척도가 개발되었다. 탐색적 및 확인적 요인분석 결과, 연구자의 가설대로 OMS의 단일차원 구조 모형이 지지가 되었고, 모형 적합도 또한 우수한 것으로 나타났다. 상담 성과로 증상 완화와 다차원적 안녕감 관점을 포함한 OMS의 단일차원 구조가 지지되어 해석적인 측면에서도 용이하고, 현장에서의 활용 가능성도 높다고 볼 수 있다. OMS 문항들은 모두 .70 이상의 요인부 하량을 지니는 것으로 나타났는데, 이는 매우 높은 요인부하량으로 간주되며, 문항 각각이 요인에 매우 잘 부합한다고 볼 수 있다. 둘째, OMS의 내적합치도는 .93으로 높은 신뢰도를 갖는 것으로 나타났다. 이는 ORS(Miller et al., 2003) 원척도의 내적합치도 .93, 국내 타당화된 ORS(안성식 등, 2020)의 내적합치도 .91과 유사한 수준이다.

셋째, OMS는 PHQ-9(우울), K-GAD-7(불안), WHO-5(안녕감) 그리고 K-CORE-OM(다영역 상담 성과)과 높은 상관을 보였다. 각각 우울과 불안 증상을 측정하는 PHQ-9 및 K-GAD-7는 주기적 결과 모니터링에서 자주 사용되어 영국 IAPT 뿐만 아니라 국내에서 현재 시행되고 있는 전국민 마음투자 지원사업의 척도로 포함되어 있는 척도이다. 안성식 등(2020)이 타당화한 ORS의 상관결과 값과 비교해 보면, ORS와 PHQ-9($r = -.54$) 및 K-GAD-7($r = .49$)와의 상관에 비해 본 연구에서 개발한 OMS는 PHQ-9($r = -.699$) 및 K-GAD-7($r = -.645$)과 더 높은 상관계수를 보였다. 이는 OMS가 ORS보다 내담자의 증상과의 관련성이 높아, 증상과 관련된 심리적 불편감을 잘 포착할 수 있다는 것을 보여주는 결과이며, OMS가 기존의 ORS가 심리적 증상을 직접적으로 탐색하지 못한 제한점을 고려하여 개발한 의의를 지지하는 결과이다. OMS는 WHO-5와도 .806의 높은 상관을 보였는데, 이는 개발한 OMS가 주관적 안녕감 및 긍정적인 정서와 높은 상관을 보여 상담 성과로 안녕감의 중요성을 강조한 관점이 반영되었음을 의미한다. 또한, OMS가 긍정적인 관점의 안녕감과 밀접한 연관성을 보이면서도 동시에 우울, 불안과 같은 증상과도 높은 관련성을 보여 상담 성과를 측정하는 균형있는 척도임을 의미한다. 마지막으로 5문항의 OMS는 상담 성과 측정 도구 중 하나로 널리 사용되는 다영역을 측정하는 34문항 K-CORE-OM과도 높은 상관($r = -.783$)을 보였는데, 이러한 결과는 초단축형 척도 OMS가 상담 성과를 타당하게 측정할 수 있는 도구임을 보여준다.

본 연구는 다음과 같은 측면에서 의의가 있다. 첫째, 매회기 혹은 회기간 자주 측정하는

것을 용이하게 하기 위해 5문항의 짧은 상담 성과 척도를 개발하였다. 문항개발 시 기존 ORS처럼 수검자에 의해 문항이 다양하게 해석되는 것을 방지하기 위해 대학생 및 상담 및 임상 전공 대학원생을 대상으로 포커스 그룹 인터뷰를 진행하여 문구를 검토하고 수정하여 문항의 의미가 분명하게 전달될 수 있도록 하였다. 둘째, 증상과 관련된 심리적 불편감을 측정함으로써 상담을 통한 증상 감소를 확인할 수 있고, 다양한 심리적 증상(예, 우울, 불안)을 호소하는 내담자들에게도 SUDS와 유사하게 우선 활용이 가능하도록 개발하였다. 다양한 내담자들에게 서로 다른 척도(예, PHQ-9, GAD-7)를 사용할 경우, 서로 다른 점수 범위를 가지고 있어 특정 검사의 점수 변화 값만으로는 내담자간 상담 경과나 성과를 비교하기가 어려울 수 있다. 하지만, OMS를 사용하는 경우 심리적 불편감 문항을 통해 증상의 변화를 추적하면서도 일반적으로 상담을 통해 회복되고 기대할 수 있는 안녕감 측면의 변화를 파악할 수 있어 서로 다른 내담자의 상담 경과를 모니터하고 성과를 파악하는 것이 가능하다. 셋째, 증상에 초점을 둔 척도들(예, PHQ-9, GAD-7)과는 달리 OMS는 자기개념, 친밀한 관계, 사회적 기능과 역할, 전반적 주관적 만족도와 같은 심리적, 사회적, 주관적 차원에서의 안녕감을 확인함으로써 반복적으로 검사를 하는 과정에서 내담자의 취약한 부분(예, 증상)만 모니터하는 것이 아니라 회복과 적응의 관점을 탐색할 수 있다는 장점이 있다. 넷째, ORS에서 채택하고 있는 시각적 아날로그 척도(Visual Analog Scale [VAS]) 방식의 문제점을 인식하여 0점부터 10점까지의 11 점 리커트 척도를 채택하였다. VAS는 채점과 해석에 오랜 시간이 소요되고, 일정 간격을

두고 반복 측정할 때는 응답자가 이전에 응답한 점수를 기억하지 못해 편향되거나 부정확하게 응답할 수 있다. 본 연구에서 개발한 OMS는 VAS의 단점을 보완해 11점 리커트 척도를 채택하였고, 응답과 채점이 모두 용이하다는 장점이 있다. 또한, 변화에 대한 민감성 측면에서 다른 5점 리커트 척도(예, WHO-5)에 비해 비교적 짧은 기간, 혹은 상담 회기간 발생하는 작은 변화를 포착할 수 있는 장점이 있다. 다섯째, 본조사에서 성별, 연령, 거주지역 별 인구비례 대비 표집을 통한 1,000명의 자료를 수집해서 진행했다는 점에서 의의가 있다. 총화표집을 통해 특정 집단이 과소 또는 과대대표되어 나타나는 편향을 줄였으며, 전체 모집단 특성을 보다 정확히 반영함으로써 연구 결과의 일반화 가능성을 높였다.

본 연구에서 개발한 OMS와 관련된 상담 실제와 상담자 훈련과 관련된 적용가능성은 다음과 같다. 첫째, 상담 서비스의 질을 향상시키기 위해 현장에서 OMS를 활용하여 주기적 결과 모니터링을 시행해 볼 수 있다. ROM이 다수의 연구에서 상담 성과에 미치는 궁정적인 효과를 고려했을 때(De Jong et al., 2021), OMS의 활용이 상담 성과를 높이는 데 도움이 될 수 있다. 구체적으로 OMS를 이용한 ROM을 통해 상담 현장에서 상담자들이 어려움을 겪는 초기 종결 및 중도 탈락을 방지하는 데 활용해 볼 필요가 있다. 둘째, OMS를 상담자 훈련 과정에서 활용해 볼 수 있다. 선행연구에 따르면, 초심상담자는 자신감이 부족하고, 내담자에게 효과적인 상담을 제공하고 있는지 자신의 상담 실력에 대해 확신을 갖기 어려워하며(김지연 등, 2009), 내담자들의 부정적인 반응이나 태도, 불만, 종결 의사 표현을 마주했을 때 어려움과 부정적인 정서를 경험한다

(김길문, 정남운, 2004). OMS의 활용은 초심 상담자들이 겪는 상담 실무에서의 불확실성을 해소하는 데 도움을 줄 수 있고, 상담 과정의 조율을 위한 치료적 대화를 촉진하는 데 활용 할 수 있다. 매회기 OMS를 실시해 내담자 변화 정도를 살펴봄으로써 자신의 상담 개입 효과성에 대해 보다 명확하게 파악할 수 있으며, 상담 성과에 대한 내담자 피드백을 반영하여 향후 어떤 방식으로 상담을 보완해 나가야 하는지 통찰을 얻는 데 도움이 될 수 있다. 또한 초심 상담자를 지도 감독하는 수퍼바이저는 OMS 자료를 통해 내담자의 관점을 확인함으로써 수퍼바이저에게 보다 객관적인 자료에 근거한 수퍼비전을 제공할 수 있다.

본 연구의 제한점 및 후속 연구를 위한 제언은 다음과 같다. 첫째, 본 연구는 대학생 및 일반 성인을 포함한 1,400명을 대상으로 인구 비례 대비 방식으로 자료를 수집한 장점이 있으나, 내담자 또는 임상 집단을 대상으로 타당화하지 못한 제한점이 있다. 앞으로 OMS가 ROM을 위한 상담 성과 측정 도구로 활용되기 위해서는 향후 상담 및 심리치료를 받고 있는 집단을 대상으로 재차 타당화하는 작업이 필요하다. 특히 추후 연구에서는 임상 집단과 비임상 집단을 구분할 수 있는 임계점을 산출하여 상담 개입의 효과 측정은 물론 임상적 결정을 내리는 데에 유용하게 활용될 수 있는지 확인해 볼 필요가 있다. 둘째, 본 연구에서는 검사-재검사 신뢰도를 파악하지 못한 한계가 있다. 따라서, 추후 검사-재검사 신뢰도를 통해 척도의 일관성과 안정성을 파악할 필요가 있는데(Nunnally & Bernstein, 1994), 특별한 개입이나 실질적인 변화가 없는 경우에는 내담자의 검사 결과가 일관되게 나오는 것이 중요하기 때문이다. 또한, OMS의 변화에 대한

민감성도 파악할 필요가 있는데, 상담이 진행됨에 따라 내담자가 상담의 효과를 통해 회복해 가는 과정에서는 회기별로 측정했을 때, OMS가 이러한 차이에 민감하게 반영하여 성과 변화를 측정할 수 있는지 탐색할 필요가 있다. 셋째, 추후 연구에서는 내담자의 실질적인 변화 및 상담 효과를 판단할 수 있는 신뢰로운 변화 지수(Reliable Change Index [RCI])를 산출할 필요가 있다. RCI는 내담자의 특정 검사 점수의 변화가 단순히 측정 오류에 의한 것인지 신뢰할 만큼 의미있는 변화인지를 결정하는 데 활용되는 통계적 지표이다(Jacobson & Truax, 1991). 따라서, 상담 전후 또는 매회기 상담 이후 OMS를 반복하는 설계를 통해 상담 개입 전후의 점수 변화를 검토하여 RCI를 산출할 필요가 있다. 넷째, OMS는 5문항으로 구성된 짧은 척도로 응답자의 부담을 줄일 수 있다는 장점이 있지만, 다른 초단축형 척도들이 지니는 제한점을 마찬가지로 지닐 수 있다. 예를 들어, 사회적 바람직성 같은 반응 편향으로 인해 내담자의 응답 결과가 실제보다 더 긍정적으로 나타날 수 있다(Miller et al., 2003). 따라서 추후 연구에서는 사회적 바람직성을 측정하는 척도와 함께 설문을 시행하여 반응 편향의 가능성은 평가해 볼 필요가 있다. 마지막으로, 후속 연구로 OMS를 활용한 ROM을 통해 상담의 효과성을 확인하는 연구를 고려해 볼 수 있다. 국외에서는 ROM 관련 연구가 다수 이루어졌고, Lambert 등(2018)의 메타분석 연구에서는 ROM 기반의 상담이 일반적인 상담보다 상담 효과가 더 좋다는 것이 밝혀졌다. 국내에서는 앞서 이화실과 이해은(2016)이 ORS를 활용한 상담과정에 대한 사례 연구를 통해 매회기 상담 성과 측정 도구를 활용하는 것이 상담 효과를 촉진할 수 있다는

예비적인 결과를 언급한 바 있다. 추후에는 국내 상담자와 내담자를 대상으로 OMS를 활용한 ROM 기반의 상담과 일반적인 상담을 비교하는 연구를 통해 상담 효과에 어떤 차이가 있는지 파악해 볼 필요가 있다. 이러한 연구를 기반으로 OMS를 활용한 상담 모니터링 및 성과 연구가 활성화된다면 상담심리 분야의 전문성이 강화되고 내담자와의 상담의 질을 향상하는 데 기여할 뿐만 아니라, 나아가 국민의 정신건강 증진에도 도움이 될 수 있을 것으로 기대된다.

참고문헌

- 고홍월 (2021). 상담의 성과는 무엇이고, 무엇으로 평가하는가? *상담학연구*, 22(5), 45-57.
- 김계현 (2000). *상담심리학 연구: 주제론과 방법론*. 학지사.
- 김광웅 (2007). 자아존중감 및 자기수용과 스트레스 대처행동과의 관계. *한국심리학회지: 상담 및 심리치료*, 19(4), 983-999.
- 김길문, 정남운 (2004). 초보 상담자가 상담 회기 내에 경험한 어려움과 대처 과정: 질적 분석. *한국심리학회지: 상담 및 심리치료*, 16(1), 1-20.
- 김선경, 왕은자 (2016). 한국판 핵심 상담 성과 도구 (The Korean Inventory of Clinical Outcomes in Routine Evaluation-Outcome Measure; K-CORE-OM)의 타당화. *상담학연구*, 17(3), 105-127.
- 김지연, 한나리, 이동귀 (2009). 초심상담자와 상담전문가가 겪는 어려움과 극복방안에 대한 개념도 연구. *상담학연구*, 10(2), 769-792.
- 김현지, 문유선, 손봉기, 이상규, 노현진, 김도훈 (2010). 지역사회에 거주하고 있는 노인의 우울증과 삶의 질을 평가하는 도구로 한국판 WHO-5의 유용성. *노인정신의학*, 14(2), 90-96.
- 문현미 (2005). 인지행동치료의 제3동향. *한국심리학회지: 상담 및 심리치료*, 17(1), 15-33.
- 보건복지부 (2024). 전국민마음투자지원사업 심리상담 표준매뉴얼. https://www.maumtuja.or.kr/course/course_view.jsp?id=4
- 안성식, 심은정, 박승민 (2020). 상담현장에서 활용가능한 초간단 척도(Ultra-brief scale)인 상담 성과척도(Outcome Rating Scale: ORS)의 국내 타당화 연구. *상담학연구*, 21(4), 27-46.
- 안제용, 서은란, 임경희, 신재현, 김정범 (2013). 한국어판 우울증 선별도구(Patient Health Questionnaire-9, PHQ-9)의 표준화 연구. *생물치료정신의학*, 19(1), 47-56.
- 오현숙, 민병배 (2008). 한국사회와 교유적 성취: 문화적 배경과 국가간 비교: 독일과 한국 청소년의 기질 및 성격 특성의 차이. *한국심리학회지: 문화 및 사회문제*, 14(1), 391-407.
- 이성숙, 김수임 (2024). 내적자원 중심 상담성과 척도(IRECOS)의 개발 및 타당화. *상담학연구*, 25(2), 207-228.
- 이화실, 김계현 (2019). 상담에서 정서 이면 욕구 확인의 치료적 효과. *상담학연구*, 20(2), 1-23.
- 이화실, 이해은 (2016). 상담성과 모니터링을 적용한 상담과정에 대한 사례연구. *상담*

- 학연구: 사례및실제, 1(2), 1-23.
- 최성수, 이영주, 박소정 (2023). 청년 1인가구의 인지된 고립감이 우울 및 주관적 안녕감에 미치는 영향. *한국웰니스학회지*, 18(4), 217-224.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed., Vol. 4). American Psychiatric Publishing.
- American Psychological Association. (2006). *Evidence-based practice in psychology: APA presidential task force on evidence-based practice*. <https://www.apa.org/practice/resources/evidence/evidence-based-statement.pdf>
- Barkham, M., Hardy, G. E., & Mellor-Clark, J. (2010). *Developing and delivering practice-based evidence*. Wiley.
- Barkham, M., McGrath, G., Leach, C., Lucock, M., Evans, C., Margison, F., Mellor-Clark, J., Audin, K. (2001). Service profiling and outcomes benchmarking using the CORE-OM: Toward practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, 69(2), 184-196.
- Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Psychology*, 3, 77-85.
- Beck, A. T. (1972). Measuring depression: The depression inventory. In T. A. Williams, M. M. Katz, & J. A. Shields (Eds.), *Recent advances in the psychobiology of the depressive illnesses* (pp. 299-302). U.S. Government Printing Office.
- Berman, P. S. (2020). *상담전문성 향상을 위한 사례개념화 원리와 실제* (이윤주 역). 학지사. (원본 출판 2015년).
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1(2), 245-276.
- Codman, E. A. (2013). The classic: A study in hospital efficiency: As demonstrated by the case report of first five years of private hospital. *Clinical Orthopaedics and Related Research®*, 471, 1778-1783.
- Cummergen, K., Hannah, L., Jopling, L., Cameron, R., Walsh, C., & Perez, J. (2022). What outcomes matter to service users who experience persistent depression: A mixed-method narrative review and synthesis. *Journal of Affective Disorders Reports*, 10, 100431.
- De Jong, K. (2016). Deriving implementation strategies for outcome monitoring feedback from theory, research and practice. *Administration and Policy in Mental Health and Mental Health Services Research*, 43(3), 292-296.
- De Jong, K., Conijn, J. M., Gallagher, R. A., Reshetnikova, A. S., Heij, M., & Lutz, M. C. (2021). Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: A multilevel meta-analysis. *Clinical Psychology Review*, 85, Article 102002.
- De Jong, K., Delgadillo, J., & Barkham, M. (2023). *Routine outcome monitoring and feedback in psychological therapies*. McGraw-Hill Education.
- Delgadillo, J., de Jong, K., Lucock, M., Lutz, W., Rubel, J., Gilbody, S., Ali, S., Aguirre, E., Appleton, M., Nevin, J., O'Hayon, H., Patel, U., Sainty, A., Spencer, P., & McMillan, D.

- (2018). Feedback-informed treatment versus usual psychological treatment for depression and anxiety: A multisite, open-label, cluster randomised controlled trial. *The Lancet Psychiatry*, 5(7), 564-572.
- Derogatis, L. R., Rickels, K., & Rock, A. F. (1976). The SCL-90 and the MMPI: A step in the validation of a new self-report scale. *The British Journal of Psychiatry*, 128(3), 280-289.
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 95(3), 542.
- Diener, E., Oishi, S., & Tay, L. (2018). Advances in subjective well-being research. *Nature Human Behaviour*, 2(4), 253-260.
- Dube, P., Kurt, K., Bair, M. J., Theobald, D., & Williams, L. S. (2010). The p4 screener: Evaluation of a brief measure for assessing potential suicide risk in 2 randomized effectiveness trials of primary care and oncology patients. *Primary Care Companion to The Journal of Clinical Psychiatry*, 12(6), PCC.10m00978.
- Ellis, A., & Harper, R. A. (1997). *A guide to rational living* (3rd ed.). Wilshire Book Company.
- Ellis, A., & Robb, H. (1994). Acceptance in rational-emotive therapy. In S. C. Hayes, N. S. Jacobson, V. M. Follette, & M. J. Dougher (Eds.), *Acceptance and change: Content and context in psychotherapy* (pp. 91-102). Context Press.
- Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry*, 180(1), 51-60.
- Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G. (2000). CORE: Clinical outcomes in routine evaluation. *Journal of Mental Health*, 9(3), 247-255.
- Eysenck, H. J. (1952). The effects of psychotherapy: an evaluation. *Journal of Consulting Psychology*, 16(5), 319-324.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). Prentice Hall.
- Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology*, 32(1), 50-55.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, 23(1), 56.
- Hofstede, G. (2001). Culture's recent consequences: Using dimension scores in theory and research. *International Journal of Cross Cultural Management*, 1(1), 11-17.
- Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T., & Stephenson, D. (2015). Loneliness and social isolation as risk factors for mortality: A meta-analytic review. *Perspectives on Psychological Science*, 10(2), 227-237.
- Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social relationships and mortality risk: A meta-analytic review. *PLoS Medicine*, 7(7), e1000316.
- Hong, S. (2020). *The applicability of the Partners for Change Outcome Management System for psychotherapy in South Korea: Exploring Korean*

- therapists' experiences [Doctoral dissertation, University of Kentucky]. UKnowledge.
- Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59(1), 12-19.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1), 141-151.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36.
- Keyes, C. L. (2006). Subjective well-being in mental health and human development research worldwide: An introduction. *Social Indicators Research*, 77(1), 1-10.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling*. Guilford Press.
- Kopta, S. M., & Lowry, J. L. (2002). Psychometric evaluation of the Behavioral Health Questionnaire-20: A brief instrument for assessing global mental health and the three phases of psychotherapy outcome. *Psychotherapy Research*, 12(4), 413-426.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2003). The Patient Health Questionnaire-2: Validity of a two-item depression screener. *Medical Care*, 41(11), 1284-1292.
- Lambert, M. J., Hansen, N. B., & Finch, A. E. (2001). Patient-focused research: using patient outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology*, 69(2), 159.
- Lambert, M. J., Vermeersch, D. A., & Brown, G. J. (2004). *Administration and scoring manual for the OQ-30.2*. American Professional Credentialing Services.
- Lambert, M. J. (2013). Outcome in psychotherapy: The past and important advances. *Psychotherapy*, 50(1), 42-51.
- Lambert, M. J., Whipple, J. L., & Kleinstäuber, M. (2018). Collecting and delivering progress feedback: A meta-analysis of routine outcome monitoring. *Psychotherapy*, 55(4), 520-537.
- Lebow, J. L. (2012). *Twenty-first century psychotherapies: Contemporary approaches to theory and practice*. John Wiley & Sons.
- Lee, S. H., Shin, C., Kim, H., Jeon, S. W., Yoon, H. K., Ko, Y. H., Pae, C. U., & Han, C. (2022). Validation of the Korean version of the Generalized Anxiety Disorder 7 self-rating scale. *Asia-Pacific Psychiatry: Official Journal of the Pacific Rim College of Psychiatrists*, 14(1), e12421.
- Lewis, C. C., Boyd, M., Puspitasari, A., Navarro, E., Howard, J., Kassab, H., Hoffman, M., Scott, K., Lyon, A., Douglas, S., Simon, G., & Kroenke, K. (2019). Implementing measurement-based care in behavioral health: A review. *JAMA Psychiatry*, 76(3), 324-335.
- Miller, S. D. (n.d.). *ORS and SRS licensing agreement*. Scott D. Miller. <https://www.scottdmiller.com/group-licensing-agreement/>
- Miller, S. D., & Duncan, B. L. (2004). *The outcome and session rating scales: Administration and*

- scoring manual. Institute for the Study of Therapeutic Change.
- Miller, S. D., Duncan, B. L., Brown, J., Sparks, J. A., & Claud, D. A. (2003). The outcome rating scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy*, 2(2), 91-100.
- Miller, S. D., Duncan, B. L., & Hubble, M. A. (2005). Outcome-informed clinical work. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (2nd ed., pp. 84 - 102). Oxford University Press.
- Nunnally, J. C. (1978). *Psychometric theory*. McGraw-Hill.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.
- Ogles, B. M. (2013). Measuring change in psychotherapy research. In M. J. Lambert (Ed.), *Bergin & Garfield's Handbook of psychotherapy and behavior change* (6th ed., pp.134-166). Wiley.
- Oslin, D. W., Hoff, R., Mignogna, J., & Resnick, S. G. (2019). Provider attitudes and experience with measurement-based mental health care in the VA implementation project. *Psychiatric Services*, 70(2), 135-138.
- Overington, L., & Ionita, G. (2012). Progress monitoring measures: A brief guide. *Canadian Psychology/Psychologie canadienne*, 53(2), 82-92.
- Pavot, W., & Diener, E. (2004). The subjective evaluation of well-being in adulthood: Findings and implications. *Ageing International*, 29(2), 113-135.
- Prins, A., Bovin, M. J., Smolenski, D. J., Marx, B. P., Kimerling, R., Jenkins-Guarnieri, M. A., Kaloupek, D. G., Schnurr, P. P., Kaiser, A. P., Leyva, Y. E., & Tiet, Q. Q. (2016). The Primary Care PTSD Screen for DSM-5 (PC-PTSD-5): Development and evaluation within a veteran primary care sample. *Journal of General Internal Medicine*, 31(10), 1206-1211.
- Rogers, C. R. (1959). A Theory of Therapy, Personality, and Interpersonal Relationships: As Developed in the Client-Centered Framework. In S. Koch (Ed.), *Psychology: A Study of a Science. Formulations of the Person and the Social Context* (Vol. 3, pp. 184-256). McGraw Hill.
- Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57(6), 1069.
- Schiepek, G., & Aichhorn, W. (2013). Real-time monitoring of psychotherapeutic change processes. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 63(1), 39-47.
- Scott, K., & Lewis, C. C. (2015). Using measurement-based care to enhance any treatment. *Cognitive and Behavioral Practice*, 22(1), 49-59.
- Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32(9), 752-760.
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, 166(10), 1092 - 1097.
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2007). The Patient Health Questionnaire Primary Care

- Study Group, & Patient Health Questionnaire Primary Care Study Group. (1999). Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *Jama*, 282(18), 1737-1744.
- Spitzer, W. O., Dobson, A. J., Hall, J., Chesterman, E., Levi, J., Shepherd, R., Battista, R. N., & Catchlove, B. R. (1981). Measuring the quality of life of cancer patients: A concise QL-index for use by physicians. *Journal of Chronic Diseases*, 34(12), 585-597.
- Topp, C. W., Østergaard, S. D., Søndergaard, S., & Bech, P. (2015). The WHO-5 Well-Being Index: A systematic review of the literature. *Psychotherapy and Psychosomatics*, 84(3), 167-176.
- Triandis, H. C. (1994). *Culture and social behavior*. McGraw-Hill.
- Wampold, B. E. (2013). *The great psychotherapy debate: Models, methods, and findings*. Routledge.
- Watson, D., Suls, J., & Haig, J. (2002). Global self-esteem in relation to structural models of personality and affectivity. *Journal of Personality and Social Psychology*, 83(1), 185-197.
- Wolpe, J. (1990). *The practice of behavior therapy* (4th ed.). Pergamon press.
- World Health Organization. (1946). Constitution of the World Health Organization. *American Journal of Public Health and the Nation's Health*, 36(11), 1315-1323.
- World Health Organization. (1988). *WHO psychiatric disability assessment schedule*.

원고 접수일 : 2024. 08. 29

수정원고접수일 : 2025. 01. 31

게재 결정일 : 2025. 02. 17

Development and Validation of the Outcome Measurement Scale for Routine Outcome Monitoring

Kyu Jin Yon¹⁾ Hyunjin Jang²⁾ A Yung Jin³⁾ Younghun Song⁴⁾
Hwajin Cho⁵⁾ Minjeong Kim⁶⁾ Chong Min Kim⁷⁾ Moonsoo Lee⁸⁾

¹⁾Sogang University, Department of Psychology / Professor

²⁾Sogang University, Department of Psychology / PhD candidate

³⁾Sogang University, Department of Psychology / M.A.

⁴⁾Sogang University, Department of Psychology / M.A. candidate

⁵⁾Hannam University, Department of Counseling Psychology / Associate Professor

⁶⁾Ajou University, Graduate School of Education / Assistant Professor

⁷⁾Gyeongin National University of Education, Department of Education / Professor

⁸⁾Gyeongsang National University, Department of Education/ Associate Professor

This study developed and validated the Outcome Measurement Scale (OMS), integrating symptom reduction and a multidimensional well-being perspective for Routine Outcome Monitoring (ROM). To address limitations of the existing Outcome Rating Scale (ORS), widely used for ROM, and based on a review of relevant literature, a brief outcome scale was developed. The OMS consists of five items measuring psychological discomfort, self-concept, intimate relationships, work and roles, and overall satisfaction. In a preliminary survey of 400 university students, item analysis and exploratory factor analysis supported a unidimensional structure. A subsequent survey of 1,000 participants from the general population confirmed this structure through confirmatory factor analysis, demonstrating excellent internal consistency ($\alpha = .93$). The OMS exhibited strong concurrent validity through significant correlations with established ROM measures, including the Clinical Outcomes in Routine Evaluation-Outcome Measure, depression (PHQ-9), anxiety (GAD-7), and well-being (WHO-5 Index). Implications for clinical practice and study limitations are discussed.

Key words : Routine Outcome Monitoring, measurement-based care, counseling outcomes, scale development