

■ 세 미 나

객관식 문항 개발 원칙

유 선 미

단국의대 가정의학교실

초 록

선택형 문항(객관식 문항)은 공정성이 필요한 의학교육의 평가에 가장 많이 사용된다. 본 강좌는 의과대학생의 교육이나 전공의 교육에 관여하는 가정의학과 의사들이 학습자를 대상으로 평가를 시행하기 위하여 선택형 문항을 개발하는데 도움이 되는 내용을 다룬다. 구체적으로 선택형 문항의 특징과 개발 원칙, 문항을 개발할 때 빠지기 쉬운 함정과 극복 방법 등에 대해 소개할 것이다.

1. 선택형 문항의 특징

객관식 문항(선택형 문항 selection-type item)은 필답 형태로 이루어지는 의학 교육의 평가에서 가장 많이 사용되는 유형이다. 학부 교육이나 전공의 과정 중에서 학습자의 학문적 성취를 측정하는 다양한 시험에서 객관식 문항이 가장 많이 이용된다. 의과대학 졸업 후 또는 전공의 과정 수련 후 치르는 자격시험 처럼 많은 사람이 동시에 치르고 공정성이 요구되는 시험에서는 현실성을 고려할 때 유일한 평가 방법이기도 하다. 객관식 문항이 많이 이용되는 것은 다음과 같은 장점을 가지고 있기 때문이다.

- 공정하고 객관적인 평가를 할 수 있다.
- 채점 업무를 자동화할 수 있다.
- 평가 대상이 되는 학습 내용과 영역을 포괄적으로 다룰 수 있다.
- 문항 분석을 위한 통계분석이 가능하다.

그러나 수험자의 개성을 표현할 수 있는 기회가 제한되고, 문항 제작에 많은 시간과 노력이 소모된다는 단점이 있다.

2. 선택형 문항의 유형

- 진위형(true-false form)
- 선다형(multiple choice form)
- 연결형(matching form)

선택형 문항의 유형 중에서 가장 널리 사용되는 것은 선다형 문항(multiple choice form)이다. 앞에 문간(stem)을 두고 밑에 제시된 4-5개의 답지(responses, options)에서 정답을 고르는 형식이다. 문항의 형식이나 정답을 고르는 방법에 따라 여러 가지 유형으로 나누어지는데, 그 중에서 A-type과 K-type이 가장 많이 쓰이고 최근 R-type이 의사 국가시험에 도입되었다.

가. A-type 문항

A-type 문항은 제시된 답지 중에서 절대적인 정답이 아니라도 가장 알맞은 답을 선택하는 최선답형

유선미: 객관식 문항 개발 원칙

(one best answer family) 문항 유형이다.

다음 질문을 읽고 가장 적당한 답 하나를 고르시오.

<지시문>

문) 10세 여아가 목감기를 앓고 약 2 주후에 맨 눈으로 구별되는 혈뇨가 나타나 병원에 왔다. 이 아이의 혈압은 170/100 mmHg 였으며, 다리 쪽에 상당한 부종이 있었다. BUN 은 3.2 mg/dL 였다. 가장 가능성 있는 원인은?

<문간>

- ㄱ) 심실 중격 결손
- ㄴ) 갑상선 기능 이상
- ㄷ) aldosterone 분비 증가
- ㄹ) catecholamine 생성 증가
- ㅁ) intravascular volume 증가

<답지>

여기에서 지시문 (direction)은 이 유형의 성격을 최선답형 (one best answer family)으로 규정해주는 중요한 것으로 빠뜨리면 수험생들이 문제를 풀 때 혼란이 생길 수 있다. 문제지의 첫 부분에 한 번 기록되거나 그에 상응하는 내용을 문간에 포함시켜야 한다.

나. K-type 문항

K-type 문항은 정답형 (true-false 혹은 absolutely correct family)에 속하는 유형으로, 답지 중에 완전한 참인 것을 골라 미리 정해진 조합에 따라 답을 하는 조합형 문항이다. 이 유형의 구조는 다음과 같다. (문항의 처음과 끝에 있는 지시문은 둘 중에 한 번만 사용해도 된다.)

다음 문제를 읽고 아래와 같은 요령으로 답하십시오.

- ㄱ) 1, 2, 3 이 맞을 경우
- ㄴ) 1, 3 이 맞을 경우
- ㄷ) 2, 4 가 맞을 경우
- ㄹ) 4 만 맞을 경우
- ㅁ) 모두 맞을 경우

<지시문>

문) 내시경 부비동 수술후 회복실에서 경과 관찰 중 periorbital swelling과 ecchymosis가 관찰되었다. 경미한 안구 운동의 장애와 안구 돌출 소견을 보였으나 시력, 시야 및 안압은 정상이었다. 이 환자에서 시행할 처치는?

<문간>

1. high dose antibiotics
2. packing removal
3. incision and drainage
4. diuretics and steroid

<답지>

- ㄱ) 1, 2, 3 ㄴ) 1, 3 ㄷ) 2, 4
- ㄹ) 4 ㅁ) 1, 2, 3, 4

(지시문)

이 유형은 복합진위형 (multiple true-false type) 이라고도 불리는데, 이것은 답지 4개가 각각이 분명한 참 혹은 거짓으로 구성된 정답형 (true-false family)에 속하는 문항이며, 이것들을 조합하여 1개 이상 다수 (multiple)의 참을 선택한다는 의미이다. 그 동안 의사 국가시험과 전문의 자격시험에 자주 사용된 유형으로, 조합형이라는 이름으로도 불린다.

그러나 이 유형은 답지에서 사실 여부를 판단하기 보다 문항 작성자의 의도를 추정하는 게임을 벌이게 되기 쉽다는 단점이 있다. 또 답지의 구성이 정답형이라는 문제가 있다. 개연성과 확률에 기반을 둔 임상 진단과 치료에 있어서 답지를 절대적인 참과 거짓으로 구별하도록 하는 것은 수험생뿐만 아니라 문항 작성자에게도 매우 어려운 일이다. 특히 문항에서 요구하는 지식 수준이 암기형 (recall)이 아닌 판단형 (interpretation)이나 문제해결형 (problem solving)이라면 아예 이론적으로 불가능할 수도 있다. 이러한 문제점 때문에 선진국의 임상의학 시험에서는 거의 사용되지 않고 국내 의사시험에서도 줄이는 추세이다.

다. R-type 문항

비교적 최근에 개발된 유형으로 최선답형 (one best

유선미: 객관식 문항 개발 원칙

answer family)에 속하며 주제, 답가지, 머리글, 문항 줄기로 구성된다.

피로를 일으키는 원인들이다.

- ㄱ) acute leukemia
- ㄴ) viral hepatitis
- ㄷ) congestive heart failure
- ㄹ) depression
- ㅁ) EB virus infection
- ㅂ) folate deficiency
- ㅅ) G6PD deficiency
- ㅇ) hereditary spherocytosis
- ㅈ) hypothyroidism
- ㅊ) iron deficiency
- ㅋ) Lyme disease
- ㅌ) hemolytic anemia
- ㅍ) miliary tuberculosis
- ㅎ) Vitamine B12 deficiency

<주제>

<답지>

아래 환자들이 호소하는 피로의 원인으로 가장 적당한 것을 고르시오

<유도문>

문) 19세 된 여자로 지난 주에 인후통과 열이 피로감과 동반되어 나타났다. 경부 임파 종대가 있었고 비장이 커져 있었다. 혈액 검사에서 백혈구 수는 5,000/mm³ 이었으며, SGOT 200 U/L를 나타내었으나 빌리루빈과 alkaline phosphatase는 정상 범위였다.

<문간>

문) 15세 여아로 ……

<문간>

문) …최근 체중 감소를 …

<문간>

이 유형은 선택형 문항의 한 유형인 연결형

(matching form)을 확장시키고 선다형(multiple choice form)으로 변형한 것으로, 흔히 확장결합형(extended matching type)이라고 불린다. 이 유형은 구조가 다소 달라 보이지만 그 내용은 A-type과 전적으로 같다. 따라서 A-type에 적용되는 모든 원칙이 똑같이 적용된다.

3. 선다형 문항 개발

문항을 개발하기 전에 우선 이 문항을 이용하여 누구를 평가할 것인지 (즉 평가의 대상) 생각해 보아야 한다. 의과대학생을 대상으로 한 평가와 전공의를 대상으로 한 평가는 그 내용과 평가의 수준이 달라야 할 것이다.

둘째, 시험의 목적을 정의해 보도록 한다. 일정한 수의 수험생을 성적순으로 선발하기 위한 시험이라면 수험생들이 순서대로 나열될 수 있게끔 성적이 배열되는 시험 문제를 만들어야 할 것이다. 그러나 전문의 자격시험처럼 일정한 수준에 도달한 수험생들에게 자격을 인정해주기 위한 시험이라면, 수험생을 성적순으로 나열하기보다는 그 수준에 도달하였는지를 평가하기에 좋은 시험 문제를 만들어야 한다.

마지막으로 문항 수, 문항의 유형과 수준, 평가하려는 내용은 혼한 문제나 응급 임상 상황과 같은 중요한 개념에 초점을 맞추어야 한다. 소위 ‘족보’에 있는 문제를 피하기 위해 별로 중요하지 않은 내용을 출제하는 것은 바람직하지 않다. 문항의 수준은 사실의 단순한 암기보다는 지식의 응용과 판단을 평가하는 것이 바람직하다. 의사 전문의 시험을 위해 문항작성을 의뢰 받은 경우라면 전공의 수련과정에서 반드시 습득해야 하는 지식과 기술을 고려하여 구체적인 문항을 개발하도록 한다(1997년 대한가정의학회 교육수련위원회에서 개발한 핵심교육과정 안을 참고).

<문간>

가. 기본 원칙

<문간>

임상의학에서 선다형 문항은 수험생의 지식과 기술 수준을 모두 평가하는데 이용된다. 즉 평가의 최종 목적은 수험생이 환자를 진료하는 능력을 측정하는데 있다. 그러므로 가능한 한 실제 임상에서 환자를 볼

때 일어나는 상황과 비슷하게 문항을 개발하는 것이 좋다. 즉 환자를 진료하는 과정에서 벌어지는 상황을 순서대로 제시하고 의사의 지식 또는 기술 수준을 암기형, 판단형, 문제해결형으로 구분하여 질문하는 것이다. 따라서 구조물의 이름을 묻거나 병명에 따라 예상되는 증상을 묻는 문제보다는 환자의 증상과 검사 소견에 따라 진단이나 진단을 위한 검사법, 치료 방법을 묻는 것이 좋다. 다음은 실제 진료상황에 따른 질문과 그와 반대되는 질문들의 몇 가지 예이다.

좋은 질문	좋지 않은 질문
· 구조물의 처리 방법 ...혈관을 어떻게 할 것인가? ...할 때 절개해야 할 것은?	· 구조물의 이름 ...은 무슨 신경인가? ...을 형성하는 fascia는?
· 진단 추정 ...증상이 있는 환자의 병명은? ...소견을 보일 때 진단은?	· 증상 유추 ...환자는 어떤 증상을? ...검사를 하면 어떤 결과가?
· 적용 치료법 ...이 경우 적절한 치료는? ...다음 단계에 시행할 검사는?	· 적응증 ...치료법의 적응증은? ...에게 하면 안 되는 치료는?

각 문항의 문간에는 내용이 명백하고 질문에 답하는데 필요한 내용이 충분히 포함되어야 한다. 답지는 가능하면 짧고 간결한 것이 좋고, 균질한 내용으로 이루어져야 한다. 문항의 내용에 정답을 암시하는 힌트가 들어있지 않아야 하며 지나치게 어려운 문제도 좋지 않다.

나. 문항 작성 요령과 예시

(1) 질문이 진료상황과 유사한 순서로 진행되도록 한다.

실제 진료 행위를 가정하고 그와 유사하게 질문을 작성하고, 그 상황에서 필요한 행동이나 적절한 판단을 물어보도록 해야 한다.

(예)
 문) 두부 손상후 내원 당시 GCS는 13점이었으나, 입원 18시간 후 GCS가 8점이 되어 CT 검사후 응급 수술을 시행하였다. 가장 흔한 원인이 되는 두개강내 출혈은?

- 가) 뇌경막하 혈종

- 나) 뇌경막상 혈종
- 다) 뇌실질내 혈종
- 리) 뇌지주막하 출혈
- 미) 뇌실질내 출혈

이 질문을 실제 상황이라고 가정해 보자. 수험생은 이미 환자의 컴퓨터 단층촬영 결과를 알고 있고 수술까지 시행한 상황이다. 그런 수험생에게 무슨 병이었는지를 물어 본다는 것은 실제 진료의 순서와 상반된다.

<수정>
 문) 두부 손상으로 내원한 환자가 내원 당시 GCS는 13점이었으나, 입원 18시간 후 GCS가 8점이 되어 CT 검사를 받았다. 이 환자에 대한 적절한 처치는?

(2) 중요한 개념, 특히 흔한 것이나 위험성이 많은 임상 상황에 초점을 맞춰야 한다.

문항을 제작할 때는 중요한, 꼭 필요한 것을 문도록 해야한다. 한정된 범위 내에서 수년간 문제를 내다보면 중복을 피하기 위해 지엽적인 내용을 묻게 되는 수가 있다. 그러나 이런 시험은 수험생들을 암기 위주의 학습으로 유도하게 된다.

(예)
 문) 정상 체온에서 enflurane의 지방/혈액 분배 계수는 얼마인가?

- 가) 27
- 나) 36
- 다) 45
- 리) 48
- 미) 60

이 문항은 전신흡입마취제인 enflurane 뿐 만 아니라 임상에서 사용되고 있는 다른 제제인 halothane, isoflurane, sevoflurane 등 모든 흡입마취제들의 지방/혈액 분배 계수를 수험생들로 하여금 외우게 하는 위험성이 있다.

(3) 문항은 가능한 한 간결하고, 그 뜻이 명백하게 표현되어 있어야 한다.

유선미: 객관식 문항 개발 원칙

문간은 필요한 내용을 충분히 설명하여야 한다. 그러나 필요 없는 말을 넣거나 쓸데없이 까다롭게 하여 수험생을 어리둥절하게 해서는 안 된다. 문항에 사용되는 문장은 잘 다듬어져야 하고 꼭 필요한 말만 골라서 간결하게 표현하고 이해하기 쉽도록 기술하여야 한다. 특히 답지는 동사를 포함하지 않는 한 개 혹은 몇 개의 단어로 가능한 한 짧게 작성하는 것이 좋다.

(예)

문) 생후 8 개월 된 여자 아기가 병원에 오기 일주일 전에 모친에 의하여 우연히 백색등공이 발견되어서 안과에 내원하였다. 환자의 왼쪽 눈에 15프리즘 정도의 내사시가 있었고 선천성 백내장이라는 진단을 받고서 수술을 받기로 하였다. 이 환자의 수술 후 시력 회복의 예후를 결정하는 중요한 요소는 무엇인가?

- ㄱ) 안구 진탕증의 유무
- ㄴ) 다른 안 질환의 동반 여부
- ㄷ) 백내장 혼탁의 상태
- ㄹ) 전신 질환의 동반 여부
- ㅁ) 전신 상태

<수정>

문) 생후 8개월 된 여아에서 왼쪽 눈에 15프리즘 정도의 내사시가 있었고 선천성 백자장 진단을 받아 수술 받기로 하였다. 이 환자의 수술 후 시력 회복의 예후를 결정하는 중요한 요소는?

(4) 용어의 정의나 개념을 묻는 문항은 문간에서 용어를 묻고 답지에 용어의 정의를 나열한다.

(예)

문) 모집단을 대표하는 표본을 반복하여 추출하였을 때 모집단 평균과 표본 평균들과의 차이들의 표준편차를 무엇이라고 하는가?

- ㄱ) 편차
- ㄴ) 표준오차
- ㄷ) 표준편차
- ㄹ) 분산
- ㅁ) 변이

<수정>

문) 표준 오차란 무엇인가?

- ㄱ) 모집단 평균과 표본 평균과의 차이의 표준편차
- ㄴ) 표본 분포의 평균들의 표준편차
- ㄷ) 모집단 평균과 표본 평균과의 차이의 평균
- ㄹ) 표집 분포의 평균과 모집단 평균과 차이의 표준편차
- ㅁ) 편차의 제곱의 평균의 제곱근

(5) 문간의 형태는 긍정문이어야 한다.

틀린 답을 찾는 것 보다 맞는 답을 찾으려 하는 것이 교육적으로 옳은 일이며, 실제 상황과 더 유사하다.

(예)

문) 다음은 대동맥판협착증의 진찰 소견이다. 이중 대동맥판협착이 심한 것을 시사하는 소견이 아닌 것은?

- ㄱ) 제 4 심음의 촉진
- ㄴ) Paradoxical S2 splitting
- ㄷ) 흉골연 우측 제2 늑간에서 진전(thrill)의 촉진
- ㄹ) Early paking of systolic ejection murmur
- ㅁ) Delayed carotid upstroke

(6) 답을 암시하는 내용이 포함되어 있지 않아야 한다.

문항 자체가 정답에 대한 단서를 주거나 암시하는 경우가 있다. 일반적으로 결코, 절대로, 반드시, 항상, 확실히, 모든 등이 들어가는 문장은 거짓인 경우가 많으며, 대개, 드물게는 등이 들어 있는 문장은 참일 가능성이 많다. 또 정답지가 오답지 보다 대개 길고 자세히 쓰여지는 경우가 많다. 답지의 위치로 정답을 추측하는 경우도 있기 때문에 정답의 위치는 가운데 물리지 않게 배치해야 한다.

(예)

문) 28세된 여자환자가 심계항진을 주소로 병원에 왔다. 청진하여 보니 midsystolic click과 late systolic murmur가 있었다. 이 환자에 대한 또 다른 청진소견으로 예견되는 것은?

- ㄱ) 일어선 자세에서는 심잡음의 크기가 아주 작아진다.
- ㄴ) Valsalva maneuver에 의하여 심잡음이 커지고 길어진다.
- ㄷ) Amyl nitrite를 흡입시키면 모든 환자에서 click

이 더 늦게 나타난다.

- ㄱ) Squatting에 의하여 click이 반드시 더 일찍 나타난다.
- ㄴ) Isometric exercise를 시키면 심잡음이 항상 더 일찍 나타난다.

- ㄱ) 2 - 4
- ㄴ) 3 - 5
- ㄷ) 4 - 6
- ㄹ) 5 - 7
- ㅁ) 6 - 8

이 문항의 답지를 보면 정답이 ㄴ 인 것을 짐작하기 쉽다 “아주”(ㄱ답지), “모든”(ㄷ답지), “반드시”(ㄹ답지), “항상”(ㅁ답지) 등의 표현이 들어 있어서 이들은 거짓임을 시사해 주고 있기 때문이다.

(예)

문1) 25세 남자가 간기능 장애와 신경증상을 주소로 입원하였다. 눈에서 관찰되는 가장 특징적인 소견은?

- ㄱ) Keiser-Fleischer ring
- ㄴ) cherry red spot
- ㄷ) cataract
- ㄹ) uveitis
- ㅁ) glaucoma

문2) 이 환자에서 ceruloplasmin치가 감소되어 있다면 혈청내 copper치는?

- ㄱ) 상승한다.
- ㄴ) 상승, 저하가 모두 일어난다.
- ㄷ) 저하한다.
- ㄹ) 무관하다.

이 증례에서는 문1에 답을 못하더라도 문2에서 힌트를 얻어 답을 쓸 수 있으므로 바람직하지 않다.

(7) 문항 답지들의 내용이 상호 독립적이어야 한다.

답지의 내용이 서로 겹치거나 한 답지가 다른 답지의 부분 집합을 이루고 있으면 정답을 알고 있는 수험생도 혼란을 일으킬 수 있으며, 때로는 정답을 추측하는 단서가 되기도 한다. “이상 모두”나 “답 없음”과 같은 답지도 이 원칙에서 벗어난 예이다.

(예)

문) Thiopental (Pentothal)의 마취유도 정주 용량 (mg/kg)은 얼마인가?

이 문항의 답지는 그 범위가 서로 겹쳐 있으며 4 mg/kg는 ㄱ, ㄴ, ㄷ의 3곳에 모두 포함될 정도이다.

(예)

문) 슬관절의 extraarticular ankylosis in extension의 원인과 치료에서 가장 중요한 구조물은?

- ㄱ) vastus lateralis
- ㄴ) vastus intermedius
- ㄷ) vastus medialis
- ㄹ) rectus femoris
- ㅁ) 이상 모두

위 문항은 K-type으로 개선하는 것이 낫다.

<수정>

(예) 문간은 같음

1) vastus lateralis
2) vastus intermedius
3) vastus medialis
4) rectus femoris

- ㄱ) 1, 2, 3 ㄴ) 1, 3 ㄷ) 2, 4
- ㄹ) 4 ㅁ) 1, 2, 3, 4

(8) 각 답지에 똑 같은 단어들(단어들이 반복되지 않도록 한다.

(예)

문) 단일양막 쌍태 임신(monoamniotic twins)에서 높은 발생빈도를 보이고 활발한 태동으로 인해 발생하며 주산기 사망률이 6%정도 되는 제대이상(cord abnormality)은 다음 중 어느 것인가?

- ㄱ) 제대의 혈종(hematoma of the cord)
- ㄴ) 제대의 협착(stricture of the cord)
- ㄷ) 제대의 낭종(cyst of the cord)
- ㄹ) 제대의 가결절(false knots of the cord)

㉑) 제대의 진결절(true knots of the cord)

(9) 답지에 어떤 논리적인 순서가 있다면 그 순서에 따라 배열한다.

(예)

문) Vogelstein이 제시한 대장암 생성과정 중 late adenoma에서 carcinoma로 이행되는데 관련된 염색체는?

- ㉒) Chromosome 5
- ㉓) Chromosome 12
- ㉔) Chromosome 18
- ㉕) Chromosome 17
- ㉖) Chromosome 13

<수정>

문) Vogelstein이 제시한 대장암 생성과정 중 late adenoma에서 carcinoma로 이행되는데 관련된 염색체는 몇 번 염색체인가?

- ㉒) 5 번
- ㉓) 12 번
- ㉔) 13 번
- ㉕) 17 번
- ㉖) 18 번

(10) 답지의 길이는 가능하면 비슷하게 하고 다소 상이할 때는 짧은 길이의 답지부터 배열한다.

(11) 답지는 균형이 있어야 하고 동질성이 있어야 한다.

(예)

문) 다음은 postpericardiotomy syndrome에 관한 설명이다. 맞는 것은 어느 것인가?

- ㉒) 개심술 환자의 10-15%에서 발생한다.
- ㉓) 원인은 불확실하다.
- ㉔) 발열, 심낭액, 혹은 흉수를 동반한다.
- ㉕) 백혈구수는 대개 $12,000/\text{mm}^3$ 이상이다.
- ㉖) 수술을 하지 않으면 더욱 불량한 예후를 보인다.

이 문항은 한 문항에서 빈도, 원인, 합병증, 진단, 치료법을 모두 다루고 있다. 이렇게 답지가 서로 실질적인 내용을 기술하고 있다면 최선답형에서 가장 참에 가까운 답지를 고르는 어렵다. 정답에 대한 이견이 있을 수 있으며 사후 논란의 대상이 될 수 있다.