

# 인지과학의 방법

— 기능적 분석 —\*

조인래\*\*

## 1. 들어가는 말

인지과학은 마음에 대한 과학이다. 인지과학자들은 지각, 사고, 기억, 언어 이해, 학습 및 여타의 정신적 현상들을 이해하고자 한다. 이러한 현상들에 대해 철학, 심리학, 인공지능, 언어학, 인류학, 신경과학 등 다양한 분야의 연구자들이 공통된 관심을 가지고 있다는 사실이 분명해지면서 일종의 학제적 연구 활동으로 발전하게 된 것이다. 그런데 이렇게 다수의 상이한 분야의 연구자들이 하나의 통합된 연구 분야를 형성하고자 하는 일은 상당히 이례적이기도 하지만, 바로 그런 이유에서 비롯하는 현실적인 어려움 역시 적지 않다. 특히, 상이한 분야의 연구자들이 방법상의 동질성을 결여한다는 사실은, 그들이 각기 제 나름의 학문적 전통 속에서 연구를 해온 점을 고려하면 자연스러운 일이지만, 통합된 학문 분야의 형성을 위해서는 주된 장애물임에 틀림없다.

이 글에서 나는 먼저 전통적으로 기능적 분석이라 불리는 방법의 내용과 성격을 규명한다. 이 과정에서 기능에 대한 두 대표적 견해, 라이트(Larry Wright)류의 견해와 커밍스(Robert Cummins)류의 견

---

\* 기초고를 읽고 유익한 논평을 해준 정인교 교수(고려대)께 감사드린

해를 부각시키고, 그들 사이의 관계가 일반적으로 논의되는 것처럼 상호 배타적인지를 묻는다. 그리고 이러한 일반적 수준의 논구를 바탕으로 하여, 나는 인지과학적 탐구 활동에서 기능적 분석의 방법이 사용되는 방식, 의의 및 한계를 해명하고자 한다.

## 2. 기능적 설명

설명은 ‘왜 어떤 사건이 발생했는가?’라는 물음에 답함으로써 그 사건의 발생에 대한 이해를 산출하려는 활동으로 흔히 이해된다. 이러한 설명 활동의 한 유력한 형태는 주어진 사건을 야기한 원인을 제시하는 것이다. 이것이 바로 통상 인과적 설명이라 불리는 것이다. “왜 성수대교는 붕괴되었는가?”는 물음에 대하여 “다리 상판 연결부의 용접이 불량했기 때문이다”라고 답하는 것은 성수대교의 붕괴를 야기한 원인을 언급함으로써 붕괴 사건을 설명하려는 전형적 사례라고 볼 수 있다.

이러한 형태의 설명과는 달리 어떤 행위의 목적을 언급함으로써 그 행위를 설명하려는 시도가 있어 왔다. 이것은 전통적으로 목적론적 설명(teleological explanation)이라 불린 것인데, 그러한 설명의 필요성을 주창한 대표적 인물은 아리스토텔레스이다. 그는 잘 알려진 것처럼 어떤 현상에 대한 완전한 이해는 그것을 야기한 네 가지 유형의 ‘원인’ 또는 설명 인자들을 밝힘으로써 성취될 수 있다고 주장하였다. 네 가지 유형의 ‘원인’이란 목적인을 포함하여 질료인, 형상인, 그리고 능동인이다. 여기서 능동인이 우리가 통상 원인이라고 부르는 것에 가장 가깝다. 결국, 아리스토텔레스는 우리가 원인이라 부르는 것을 지목하는 것만으로 어떤 현상에 대한 완전한 설명이 이루어질 수 없다고 생각한 것이다. 가령, 의자라는 사물에 대한 이해는 무엇이 그것을 산출했는가를 아는 것만으로 충분하지 않으며, 그것을 구성하는 물질, 그것의 구조(또는 형상), 나아가서 그것의 목적 내지

기능을 파악할 때 비로소 완전하다는 것이 아리스토텔레스의 생각이었다고 말할 수 있다. 즉, 의자에 대한 이해는 그것이 목수에 의해 만들어졌다는 사실을 아는 것만으로 충분하지 않으며, 의자가 나무로 만들어졌다는 사실, 일정한 넓이의 나무판과 그것을 떠받치는 일정한 길이의 다리들로 이루어진 구조를 가진다는 사실, 나아가서 의자가 수행하는 기능을 알 때 완전해진다는 것이다. 설명에 대한 이와 같은 견해는 비단 의자와 같은 인공물에 대해서만 아니라 돌맹이와 같은 자연물에 대해서도 마찬가지로 적용된다. 예를 들어, 어떤 사람이 던진 돌맹이의 포물선 운동은 던진 사람이 돌맹이에 가한 물리적 힘, 즉 능동인에 의한 설명만이 아니라, 돌맹이를 구성하는 질료 및 그 본성과 연계되어 있는 목적인, 즉 자연적 위치(natural place)를 찾아가는 운동에 의한 설명을 필요로 하는 것으로 간주된다. 결국, 아리스토텔레스에 의하면, 네 가지 유형의 '원인'들 중의 하나에 의한 설명은 일반적으로 그 자체로는 설명의 대상에 대해 부분적인 이해만을 산출하며, 네 부류의 설명이 합쳐질 때 완전한 이해를 산출한다는 점에서 이 설명들은 대립적인 관계가 아니라 상호 보완적인 관계에 있다.

그러나, 목적론적 설명을 모든 사물들에 보편적으로 적용하는 아리스토텔레스의 전략은 내재적으로 문제를 안고 있는 것처럼 보인다. 인공물이 아닌 대상들에 목적이나 기능의 개념을 적용하는 것이 어떤 의미를 가지는가 라는 물음이 제기될 수 있기 때문이다. 특히 목적론적 설명의 보편적 적용은 일종의 인간중심주의(anthropomorphism)를 전제로 할 때에만 정당화될 수 있다는 지적을 면하기 어렵다. 실제로 아리스토텔레스의 표준적인 설명 이론은 근대의 과학혁명을 겪으면서 심각한 도전에 직면하게 된다. 즉, 근대 과학혁명의 중심체인 뉴턴 역학은 목적이나 기능 개념의 사용 없이 돌맹이나 운석 같은 자연물의 운동을 설명하는 것을 가능하게 해주었다.<sup>1)</sup> 그렇다면, 동일한 비목적론적 설명 방식이 생명체 및 생물 현

1) 참조: Sober(1993), 83.

상에도 적용될 수 있지 않을까?

### 3. 기능에 대한 분석

#### 3.1 험펠의 분석

험펠(Carl Hempel)은 과학적 설명에 대해 체계적인 이론을 개발하고자 했던 현대의 대표적 철학자이다. 그의 설명 이론은 크게 세 가지의 모형으로 구성되는데, 그중 하나가 연역·법칙적 설명 모형(deductive-nomological model of scientific explanation)이다. 이 모형에 따르면, 과학적 설명이란 특정한 조건들을 만족시키는 연역 논증이다. 이 조건들 중의 하나에 의하면 과학적 설명은 적어도 하나 이상의 법칙적 진술들을 포함해야 하며, 연역 법칙적 설명이란 이러한 법칙적 진술들을 포함하는 설명항의 진술들로부터 설명 대상인 사건이나 사태를 기술하는 피설명항의 진술을 연역적으로 도출하는 과정이다. 예를 들어, 수면 위의 공기 속을 통과하던 빛이 물 속으로 진입하면서 특정한 각도로 굴절하게 되는 현상에 대한 연역 법칙적 설명은 입사각과 굴절각 사이의 규칙적 상관 관계를 규정하는 스넬의 법칙과 주어진 상황에서의 특정한 입사각에 대한 진술로부터 피설명항을 구성하는 굴절각에 대한 진술을 연역적으로 도출하는 과정이다.

설명항이 법칙적 진술을 포함해야 한다는 조건 외에도 설명항의 진술들이 피설명항의 진술을 연역적으로 함축해야 한다는 조건, 설명항에 포함된 법칙적 진술들이 피설명항의 진술을 연역적으로 도출하는 과정에서 실질적으로 사용되어야 한다는 조건, 설명항을 구성하는 진술들이 경험적으로 시험가능 해야 한다는 조건 등이 연역법칙적 설명의 성립 요건들이다. 실제로, 빛의 굴절 현상에 대한 앞에서 언급된 설명은 이러한 조건들을 모두 만족시키는 설명의 사례이다. 그

런데 험펠의 설명 이론은 물리적 현상들에 대한 설명뿐만 아니라 여타의 모든 경험과학적 설명들을 적용 대상으로 삼고자 하는 이론이다. 그렇다면 생물학이나 사회 과학들에서 자주 등장하는 기능적 설명은 어떻게 이해되어야 하나?

먼저 다음과 같은 기능 부여 진술을 생각해 보자.

- (1) 척추동물에서 심장 박동의 기능은 그 유기체 내의 혈액을 순환시키는 것이다.

여기서, 혈액 순환은 심장 박동의 한 결과이다. 기능 부여 진술에 대한 분석에서 해결해야 할 주요 과제는, 이 결과와 심장 박동의 다른 결과들(예를 들어, 심장 박동 소리)을 차별화하는 것이다. 이를 위해 (1)을 다음과 같이 분석해 보자.

- (2) 척추동물에서 심장 박동은 혈액을 순환시키는 효과를 가진다. 그리고 이를 통해 그 유기체가 적절하게 작동하는데 필수적인 조건들(즉, 영양의 공급과 노폐물의 제거)을 만족시키도록 한다.

그런데, 험펠(1965, 305-6)에 따르면, (2)와 같은 식으로 이해된 기능 진술은 어떤 설명력도 가지지 못한다. 그 이유는, 심장이 문체의 기능을 수행하기 위해 충족되어야 할 조건들이 명시되지 않았을 뿐만 아니라 유기체의 적절한 작동(또는 정상적 기능)에 대한 기준들이 제시되지 않았다는데 있다.

대안으로, 기능에 대한 다음과 같은 분석을 생각해 보자.<sup>2)</sup>

- (3) (a) 시점  $t$ 에,  $s$ 는 일정한 내적 및 외적 조건들이 충족되는 환경  $c$  속에서 적절하게 기능한다.

---

2) Hempel(1965), 310.

- (b)  $s$ 는 어떤 필요 조건  $n$ 이 충족될 때에만  $c$  속에서 적절하게 기능한다.
- (c) 특질  $i$ 가  $s$  속에 존재한다면, 그 결과로 조건  $n$ 이 충족될 것이다.
- (d) 그러므로, 시점  $t$ 에, 특질  $i$ 는  $s$  속에 존재한다.

(3)을 있는 그대로 받아들인다면, 그것 역시 설명력을 가지지 못한다. 왜냐하면, 설명항을 구성하는 (a), (b), 그리고 (c)로부터 피설명항에 해당하는 (d)가 연역적으로 도출되지 않기 때문이다. 이 문제점을 해결하기 위해 (c)를 (c')(즉, 조건  $n$ 은 특질  $i$ 가  $s$  속에 존재할 때에만 충족될 것이다.)로 대체하면, 수정된 분석은 설명력을 가지겠지만 그릇된 전제(즉, (c'))에 의존하기 때문에 받아들여질 수 없다. 결국, 햄펠에 의하면, '기능적 설명'은 연역적 도출의 조건을 만족시키지 못하기 때문에 설명의 자격을 갖추지 못하거나, 문제의 조건을 만족시키는 방식으로 수정하면 그릇된 설명이 되는 딜레마에 빠지게 되어, 과학적 설명으로서의 적합성을 결여하는 경우이다.

### 3.2 네이글의 분석

햄펠과 같은 논리경험주의자들이 주창한 과학철학에 대한 주요 동조자인 네이글(Ernest Nagel)은 기능 부여 진술에 대한 다음과 같은 분석을 하나의 해결책으로 제안한다.<sup>3)</sup>

- (4) 조직  $C$ 를 가진 체계  $S$  속에서  $A$ 의 기능이 환경  $E$  속에서  $S$ 가 과정  $P$ 에 관여할 수 있게 하는 것이라는 진술은 다음과 같은 조건들이 성립함을 의미한다. 즉,
  - (a) 조직  $C$ 를 가진 모든 체계는 환경  $E$  속에서 과정  $P$ 에 관여한다.

---

3) Nagel(1961), 403.

- (b) 조직 C를 가진 체계 S가 환경 E 속에서 A를 가지지 않는다면, S는 P에 관여하지 않는다.
- (c) 그러므로, 조직 C를 가진 S는 A를 가져야 한다.

네이글의 해결책은 사실상 (3)에서 (c)를 (c')으로 교체한 것과 다르지 않다. 그렇다면, (c')을 어떻게 정당화할 것인가 하는 문제는 여전히 남아 있다. 이에 대한 네이글의 대응책은 다음과 같다.

- (5) 심장의 존재가 모든 상황에서 혈액 순환을 위해 필수적이지는 않을지라도 정상적(또는 대부분의) 상황에서는 그렇다.

문제는 (c')이 참이 되는 문제의 정상적 상황들을 어떻게 규정할 것인가 여부이다. 네이글의 제안에 대해 문젯거리가 되는 사례를 드는 것은 그렇게 어렵지 않다. 예를 들어, 인간의 한 쪽 신장은 정상적 상황에서 노폐물을 제거하는 기능을 가진다. 그러나 그것의 존재가 필수적이지는 않다. 왜냐하면 다른 쪽 신장이 그 역할을 담당할 수 있기 때문이다.

### 3.3 라이트의 해법

기능에 대한 분석과 관련하여 험펠이 부각시킨 딜레마를 극복하는 한 가지 방안이 라이트(Larry Wright)에 의해 제시되었다. 라이트의 해결책은 설명의 적합성을 위해 험펠이 제시한 조건, 특히 설명항이 피설명항을 연역적으로 함축해야 한다는 조건을 거부하는 것이다. 이러한 해결 방안의 이론적 원천은 진화론이다.

앞서 언급한 대로 뉴턴 역학은 물체들의 운동을 설명하기 위해 목적이나 기능에 호소해야 할 필요성을 제거하였다. 그리고 뉴턴 역학은, 그것의 획기적 성공 덕분에, 근대의 학문적 활동을 위한 전범 역할을 하였다. 이런 상황에서 생명 현상의 설명 역시 목적이나 기능에

호소하지 않고 이루어질 수 있을 것이라는 기대가 생겨났다면, 이는 하등 이상한 일이 아니다. 실제로 19세기 중반에 다윈은 종의 기원을 설명하는 과정에서 신의 역할을 제거함으로써 결과적으로 과학적 유물론의 위상을 강화하는데 결정적인 기여를 하였다. 그러나 그가 한 일은 목적이나 기능에 의존하지 않고 종의 기원을 설명하는 것과는 거리가 멀었다. 다윈의 진화론에서 종의 유지나 분화는 개체가 자손들을 번식하는 과정에서 무작위적으로 발생하는 변이와 각 개체의 표현 형질들에 대한 자연 선택의 기제를 통해 일어난다. 결국, 기존의 창조론에서 신이 하던 역할을 다윈의 진화론에서는 자연이 수행하게 되는 중요한 변화가 일어난다. 그러나, 다윈은 이 과정에서 기능과 같은 목적론적 개념을 제거하기보다는 그것이 자연주의적 틀 속에서 어떻게 유지되고 사용될 수 있는가를 보여 주었다.<sup>4)</sup>

기능에 대한 라이트의 분석은 이러한 진화론을 배경으로 하여 나온 것으로 볼 수 있다. 먼저, 그는 어떤 특질의 기능을 언급하는 것이 그 특질이 어떻게 존재하게 되었는가를 설명한다고 말한다.<sup>5)</sup> 즉, 라이트는 기능적 설명의 대상이 어떤 기능을 가지는 사물이나 특질의 존재라는 생각을 험펠과 공유한다. 이 점은 기능에 대한 그의 다음과 같은 분석에서 잘 드러난다.

- (6) X의 기능이 F라는 진술은 다음의 (a)와 (b)를 의미한다. 즉,  
 (a) X는 그것이 F를 행하기 때문에 존재하며, (b) F는 X가 존재한다는 사실의 결과이다.<sup>6)</sup>

이 분석에 따르면, “엽록소의 기능은 광합성을 하는 것이다”는 주장이 의미하는 것은 “엽록소는 광합성을 하기 때문에 존재하며, 광합성은 엽록소의 존재로부터 비롯된 것이다”가 된다. 그러나, 기능적

4) 참조: Sober(1993), 83.

5) 참조: Wright(1973).

6) Wright(1973), 42.



설명의 적합성에 대한 라이트의 견해는 험펠의 견해와 중요한 점에서 다르다. 험펠은 기능적 설명 역시 다른 설명과 마찬가지로 그가 제시한 조건들을 만족시켜야 하는 것으로 본다. 즉, 기능적 설명에서도 설명항은 피설명항을 연역적으로 함축해야 한다는 것이 그의 생각이다. 그러나 라이트는 이러한 요구 조건을 거부한다. “X는 F를 행하기 때문에 존재한다”는 주장이 “X의 기능은 F이다”는 주장의 일부를 구성한다고 제안할 때, 라이트는 설명항에 의한 피설명항(즉, X의 존재)의 연역적 함축 같은 것을 염두에 두고 있지 않다. 유기체의 특질 X를 대상으로 하여 “X는 F를 행하기 때문에 존재한다”는 주장을 할 때 라이트가 염두에 두고 있는 것은, 특질 X가 F를 행하는 것이 X를 가지는 개체나 그 후손의 생존을 위해 유리하고 따라서 그 개체나 후손이 자연에 의해 선택됨으로써 X를 가지는 개체들이 생존하게 된 상황이다.<sup>7)</sup> 이러한 상황에서 “X가 F를 행한다”는 사실이 “X의 존재”를 보장하지 못한다. 가령, X와는 다른 Y도 F를 행할 수 있다고 하자. 이러한 가상적 상황에서, Y를 가진 개체가 생존하고 번식할 확률은 일반적으로 X를 가진 개체가 생존하고 번식할 확률 못지 않을 것이며, 따라서 X의 존재는 보장되지 않을 것이다. 이처럼 “X가 F를 행한다”는 사실이 “X의 존재”를 야기한다는 것은 맥락적인 요인들에 의해 크게 좌우되는 사건이다. 라이트의 관점에서 볼 때 중요한 것은, “X가 F를 행한다”는 사건과 “X의 존재”라는 사건 사이에 실제로 인과적 연결이 존재하는가 문제일 것이다. 마찬가지로, 기능에 대한 라이트의 분석에서, (b)의 주장 역시 “X의 존재가 F를 보장한다”거나 “F가 성립할 때마다 X가 존재한다”는 것을 의미하지 않는다. 그것은 “X의 존재가 F의 성립을 위해 필요한 상황들이 있다”는 것을 의미한다.<sup>8)</sup>

‘선택된 효과’로서의 기능 개념을 제안하면서 라이트가 강조하는

7) 참조: Wright(1973), 159, 162-4. Neander(1991, 173-4)는 이러한 진화론적 고려를 명시적으로 포함하는 생물적 기능에 대한 분석을 제시한다.

8) 참조: Wright(1973), 41.

것은, 기능에 대한 어떤 분석이 적합하려면 기능에 해당하는 효과와 그렇지 못한 효과를 구분할 수 있는 기준을 제시할 수 있어야 한다는 것이다.<sup>9)</sup> 물론, 그는 그 자신의 기능 개념을 채택할 경우 이 문제는 만족스럽게 해결된다고 생각한다.<sup>10)</sup> 다음의 예를 생각해 보자. 척추동물에서 심장 박동은 그 유기체 내의 혈액을 순환시키는 효과를 낳는 동시에 박동의 소리를 내는 효과를 낳는다. 이 사례에서 우리는 혈액 순환을 심장 박동의 기능이라 부르지만, 박동의 소리를 심장 박동의 기능이라 부르기를 원하지 않는다. 기능이라는 용어의 이러한 선택적 사용이 정당하다면, 기능에 대한 적합한 분석은 이러한 용법을 정당화해 줄 수 있어야 할 것이다. 라이트의 기능 개념을 채택하면 이 문제는 어떻게 해결 가능한가? 인간의 경우, 심장 박동에 의해 산출되는 혈액 순환은 영양분의 공급과 노폐물의 제거를 위해 필요하며, 후자의 일들은 인간 개체의 생존을 위해 필수적이다. 따라서, 심장은 혈액을 내보내는 일을 하기 때문에 선택되었다. 반면, 심장 박동이 만들어내는 소리는 인간의 생존을 위해 필수적이지 않으며, 다른 이유로 인해 선택된 심장과 그 박동의 부수적 산물에 불과하다. 이처럼 선택된 효과로서의 기능 개념을 적용하면, 기능으로서의 효과와 부수적인 효과 사이의 구분이 별 무리 없이 이루어질 수 있는 것처럼 보인다.

그러나 선택된 효과로서의 기능 개념은 그 나름의 문제점이 없지 않다. 먼저, 과거의 생물학자들은 진화론에 대해 들어본 적이 없는 상황에서도 기관들에 기능을 부여하였다. 예를 들어, 하비는 진화론이 등장하기 훨씬 전인 17세기에 이미 심장의 기능은 혈액을 내보내는 것이라는 주장을 하였다. 하비가 심장 박동이 심장의 기능이라고 말했을 때, 그는 인간이 심장을 가지는 이유에 대해서라기보다는 심장이 하는 일에 대해 어떤 주장을 한 것으로 볼 수 있다. 이는 기능 개념이 선택된 효과라는 의미와는 다른 의미로 사용되었음을 시사한

9) 참조: Wright(1973), 141-2.

10) 참조: Wright(1973), 165-6.

다.<sup>11)</sup> 물론 기능 개념의 모든 용법이 동등한 지위를 가진다는 주장을 하려는 것은 아니다. 그러나 기능 개념의 다른 용법들이 존재할 경우, 각 용법의 근거와 그 정당성에 대한 검토가 필요하다. 이 외에도 라이트의 분석은 이상한 결과들을 낳는다는 지적이 있어 왔다. 예를 들어, 비만으로 인해 운동을 하지 못하는 어떤 남자의 경우를 생각해 보자. 그가 계속 비만인 상태에 있는 것은 운동을 하지 못하기 때문이기도 하다. 그렇지만 운동을 방해하는 것이 비만의 기능이라고 말하는 것은 어불성설이다.<sup>12)</sup> 이러한 사례는 라이트의 조건들을 만족시킴에도 불구하고 기능으로 간주하기 어려운 경우들의 존재를 시사한다. 그러나 라이트가 제시한 기능의 조건들이 적합하지 않다는 주장과 선택된 효과로서의 기능 개념이 잘못되었다는 주장은 구분될 필요가 있다. 라이트가 제시한 기능의 조건들을 만족시킴에도 불구하고 기능으로 간주되기 어려운 사례들의 존재는 전자의 주장을 지지하는 반대 사례의 역할을 할지언정, 후자의 주장을 위해서도 동일한 효과를 가지는 것은 아니다. 실제로 비만이나 운동 방해는 진화의 과정에서 선택된 특질이나 효과와는 거리가 멀다. 따라서, 그것이 라이트의 조건들을 만족시킨다고 하더라도 선택된 효과로서의 기능 개념에 대한 직접적인 반대 사례로 보기 어렵다.<sup>13)</sup>

### 3.4 커밍스의 대안

#### 3.4.1 기능에 의한 분석

기능적 설명의 딜레마를 극복하는 커밍스(Robert Cummins)의 방안은 피설명항의 설정이, 즉 기능을 가진 사물의 존재가 설명되어야 한다는 생각이 잘못되었다고 말하는 것이다. 어떤 것이 하는 일(즉, 기능)에 의해 그것의 존재를 설명할 수 있는 경우들이 없지 않다는

11) 참조: Boorse(1976), 74.

12) Boorse(1976), 75-6.

13) Boorse(1976, 75-6) 자신도 이 점을 인지한다.

것을 커밍스 역시 인정한다. 그러나, 그에 따르면, 이러한 경우는 인조물의 경우에 한정되며, 유기체의 경우 사정이 다르다.<sup>14)</sup> 유기체 O가 어떤 구조 S를 포함하는가는 S의 기능과는 무관하게 O가 물려받은 그리고 종 O에 특징적인 유전적 “설계”에 의해 결정되며, 따라서 구조 S의 기능을 규정하는 것이 그 구조의 존재를 설명해 주지는 않는다는 것이다.<sup>15)</sup> 그렇다면, 우리는 왜 어떤 것이 하는 일(즉, 기능)로부터 그것의 존재를 추론해 내는 방식으로 설명을 성취하려고 시도해 왔는가? 커밍스의 진단에 따르면, 이러한 시도는 설명과 설명으로의 추론을 혼동하는 데서 비롯한다. 즉, 우리는 자주 어떤 기능의 수행으로부터 어떤 특정한 구조나 과정으로의 추론을 행한다. 예를 들어, 광합성으로부터 엽록소의 존재 또는 광합성이 이루어지는 과정의 존재를 추론한다. 그러나, 이러한 추론은 최선의 설명으로의 추론에 해당하는 것이지 그 자체가 설명은 아니다. 이러한 진단이 옳다면, 잘못된 것은 험펠, 네이글 등이 함께 받아들였던 가정, 즉 과학에서 기능 진술의 의의는 어떤 기능을 가지는 것으로 간주되는 기관, 구조 또는 과정의 존재를 설명하는데 있다는 가정이었다.

그렇다면, 무엇이 기능적 설명의 대상이어야 하는가? 커밍스(1975, 54)에 의하면, 어떤 것이 하는 일에 의해 우리가 설명하거나 설명할 수 있는 것은 그것을 포함하는 계의 행동이다. 이러한 견해는 기능과 성향(disposition) 사이의 관계에 대한 논의에서 비롯한다. 어떤 계 s 속에서 어떤 구조 x의 기능이 F를 행하는 것이라면, 그 x는 s 속에서 F를 행할 수 있어야 한다. 이처럼 기능을 부여하는 진술은 성향을 부여하는 진술을 함축한다. 즉, s 속에서 x의 기능이 F를 행하는 것이라면, 그 x는 s 속에서 F를 행하는 성향을 가져야 한다. 다른 한편, 어떤 대상 a에게 성향 d를 부여하는 것은 그 a의 행동이 어떤 법칙적 규칙성을 드러낸다고 주장하는 것이다. 그리고 이 성향적 규칙성은 어떤 종류의 대상에 특수한 규칙성이며 그러한 종류의 대상

14) 참조: Cummins(1975), 53 & 56.

15) 참조: Cummins(1975), 56.

에 대한 어떤 특수한 사실들에 의해 성립하는 것이다. 따라서, 성향의 부여는 설명을 요구한다.<sup>16)</sup>

그런데 성향에 대한 설명은 그것이 사례화되는 방식에 대한 설명을 요구한다. 왜냐하면, 어떤 성향이 사례화되는 방식을 이해할 때, 우리는 그것의 규칙적 발현을 이해할 수 있는 입장에 놓이기 때문이다. 예를 들어, 물에 녹는 성향은 어떤 종류의 분자적 구조로서 사례화된다. 그러나 단순하지 않은 성향에 대한 설명은 일단 그것을 다른 성향들로 분석하는 절차를 거칠 필요가 있다. 그것이 사례화되는 방식을 바로 설명하기가 여의치 않기 때문이다. 여기서 커밍스(1975, 63)의 제안대로 성향 대신 역량(capacities)에 대해 논의한다면, 방금 언급된 분석적 절차는 더 자연스럽게 이해될 수 있다. 예를 들어, 계산기의 계산 능력은 그것을 구성하는 부품들의 보다 단순한 정보처리 능력들에 토대를 두고 있다. 따라서, 계산기의 특정한 계산 능력은 부품들의 갖가지 정보처리 능력들 및 그 조직화로 분석될 수 있다. 커밍스(1975, 63)에 따르면, 이러한 분석적 접근은 유기체의 생물학적 역량을 설명하는데 사용될 수 있다. 즉, 유기체의 생물학적 역량은 그 유기체를 구성하는 기관들의 갖가지 역량들 및 그 조직화에 의해 설명 가능하다는 것이 그의 생각이다. 그리고 이러한 생각을 바탕으로 커밍스(1975, 64)는 어떤 특질 X의 기능 F를 다음과 같이 분석한다. 즉,

- (7) 어떤 체계 s가 어떤 행위 G를 행할 역량을 가진다는 사실에 대한 분석적 설명 A와 상관하여, (i) 어떤 특질 x가 s 속에서 F를 행할 수 있으며 (ii) A가 부분적으로는 s 속에서 x가 F를 행하는 역량에 의해 G를 행하는 s의 역량을 적절하고 적합하게 설명할 때에 그리고 그럴 때에만 s 속에서 x의 기능은 F를 행하는 것이다.

그런데, 이러한 분석적 설명의 의의는 한결같지 않다. 커밍스(1975,

---

16) 참조: Cummins(1975), 61.

66)는 분석적 설명의 의의를 헤아리는데 사용될 수 있는 세 가지 기준을 제시한다. 첫째, 분석적 설명의 의의는 분석하는 역량들이 분석되는 역량보다 단순하면 할수록 크다. 둘째, 분석하는 역량들과 분석되는 역량들의 유형상 차이(difference in type)가 클수록 분석적 설명의 의의 역시 크다. 셋째, 주어진 체계를 구성하는 요소들의 조직화가 복잡할수록 관련된 분석적 설명의 의의가 크다. 여기서 세 번째 기준은 첫 번째 또는 두 번째 기준과 상관 관계가 있다. 두 부류의 역량들 사이에 존재하는 고도화 및 유형의 차이가 클수록, 그 차이를 메우기 위해 더 복잡한 조직화가 요구될 것이기 때문이다.

#### 3.4.2 기능적 분석의 유형

기능에 의한 역량의 분석은 한 가지 형태로만 이루어지는 것이 아니다. 호그랜드(John Haugeland)는 형태적(morphological) 분석과 체계적(systematic) 분석을 구분한다.<sup>17)</sup> 이러한 구분의 필요성은 분석하는 기능들 사이의 상호작용이 기능적 분석의 설명력을 결정하는데 중요한 역할을 하는 경우들과 그렇지 않은 경우들이 존재한다는 데서 비롯한다. 구체적인 사례들을 생각해보자. 먼저 광섬유들의 뭉치로 만들어진 코드의 한 쪽 끝으로부터 다른 쪽 끝으로 어떤 이미지가 전송되는 경우에 대해 설명을 한다고 하자. 여기서 설명 대상은 광섬유들의 뭉치가 지닌 이미지 전송 능력이다. 그리고 설명은 두 가지 기본적 사실을 토대로 하여 이루어질 것이다. 하나는 주어진 코드의 구성 요소들(즉, 광섬유들)이 어떤 특정한 구조(즉, 서로 평행하다는 구조)를 가진다는 사실이다. 다른 하나는 문제의 구조를 형성하는 구성 요소들(즉, 광섬유들)이 각각 특정한 역량(즉, 세지 않게 빛을 전송하는 능력)을 가진다는 사실이다. 나아가서, 설명은 문제의 역량을 가지는 요소들이 언급된 바의 구조를 가짐으로써 형성되는 모든 체계는 설명 대상에 해당하는 역량을 가진다는 주장을 포함할 것이다. 호그랜드는 이러한 유형의 설명을 형태적 분석이라 부른다. 다른

17) 참조: Haugeland(1978), 216.

한편, 자동차 엔진이 작동하는 방식을 설명하는 경우를 생각해 보자. 이 경우 설명은, 어떤 체계를 구성하는 요소들이 형성하는 구조 및 그들이 가진 역량에 덧붙여, 구성 요소들이 조직적으로 상호 작용하는 양식을 토대로 하여 이루어질 것이다. 즉, 자동차 엔진의 차를 움직이는 역량은 엔진을 구성하는 부품들이 그들 나름의 갖가지 역량과 그들 사이의 특정한 구조를 바탕으로 하여 체계적으로 상호 작용하는 데서 비롯하는 것으로 설명될 수 있다. 따라서, 이는 호그랜드가 체계적 설명이라 부르는 것의 한 예가 된다. 그런데 체계들이 기능적 요소들에 해당하는 더 작은 체계들로 구성될 뿐만 아니라 후자의 하부 체계들 역시 그보다 더 작은 체계들로 구성되는, 즉 체계 내의 체계들의 계층이 성립하는 일이 자주 일어난다. 이럴 경우, 계층적 질서의 각 수준에서 체계적 설명이 요구되기도 하는데, 이를 호그랜드(1978, 251)는 체계적 환원(systematic reduction)이라 부른다. 그리고 체계 내 계층적 질서의 가장 낮은 수준에서야 체계적 설명과는 다른 종류의 설명, 전형적으로 형태적 설명이 등장한다는 것이 그의 생각이다.

형태적 설명과 체계적 설명의 구분에 덧붙여, 커밍스(1983, 32)는 해석적/서술적 구분(interpretive/descriptive distinction)을 제안한다. 이 후자의 구분은 기능적 분석의 과정에서 입력과 출력들을 단순히 서술하기보다 해석하는 방식으로 분석하는 기능들(analyzing functions)을 규정하게 되는 상황에서 생겨난다. 호그랜드(1978, 262)에 따르면, 앞서 언급된 체계적 환원의 과정은 하나의 “동일한” 해석 하에서 수행된다. 그런데 이 체계적 환원이 가장 낮은 수준에 이르면, 기존의 해석을 그만두지 않고서는 더 이상 설명이 진행될 수 없게 된다. 그리고 기존의 해석을 그만둔 뒤에 이루어지는 설명은 사례화에 의한 설명(explanation by instantiation)의 형태를 띠게 된다. 컴퓨터의 경우를 생각해 보자. 컴퓨터는 우리가 정보 처리 체계(information processing system)라 부르는 것의 전형적인 경우이다. 그리고 정보 처리 체계로서의 컴퓨터가 지니는 여러 가지 역량들은

컴퓨터를 구성하는 여러 하부 요소들의 역량들 및 그들 사이의 조직화에 의해 설명될 수 있다. 나아가서, 각 하부 요소의 역량들에 대한 설명은 그 하부 요소를 구성하는 또 다른 하부 요소들의 역량들 및 그 조직화에 의해 가능하다. 이러한 점에서 컴퓨터의 체계 내적 하부 요소들은 그 자체 정보 처리 체계에 해당한다. 그런데 이러한 분석의 과정을 거쳐, 우리는 “연접 게이트”와 같은 대상으로 구성되는 수준에 이르게 된다. 이 연접 게이트는 호그랜드가 “단순한 의도적 검은 상자(simple intentional black box)”라 부르는 것에 해당한다. 여기서 단순하다는 것은 기존의 해석을 유지하면서 주어진 체계의 역량을 그 하부 요소들의 역량과 그 조직화에 의해 설명할 수 있는 그러한 하부 요소들이 존재하지 않는다는 의미이다. 따라서 문제의 체계는 정보 처리 체계로서의 지위도 상실한다. 그러면, 연접 게이트와 같은 대상의 역량은 어떻게 설명할 것인가? 가능한 방안은 그것의 역량을 전기 회로 및 그 구성 요소들에 의해 설명하는 것이다. 그러나 이러한 형태의 설명은 기존의 해석을 포기할 때 가능하며, 사례화에 의한 설명에 해당하는 것이다. 물론, 연접 게이트의 경우, 그것은 물리적 사례화라고 말할 수 있다. 왜냐하면, 사례화 설명에 등장하는 구성 요소들이 물리적 용어로 표현되기 때문이다. 그리고 이 사례화 설명은 기존의 해석을 철회한 가운데 이루어진다는 점에서 커밍스가 서술적 설명-해석적 설명과 대비되는 의미에서-이라 부르는 것에 해당한다고 볼 수 있다.

지금까지 논의된 기능적 분석의 유형들이 모든 가능한 경우들을 포섭할 것이라고 주장하지는 않겠다. 그러나, 이러한 분류는 기능적 분석의 많은 경우들에 대해 그 성격을 이해하는데 도움이 될 뿐만 아니라, 새로운 상황에서 요구되는 기능적 분석의 형태를 선택하거나 고안하는데 안내자 역할을 할 것이다.



#### 4. 두 대안 사이의 관계 : 상호 배타적인가?

앞 절에서 우리는 기능적 분석에 대해 비교적 최근에 제시된 두 가지 유력한 견해를 살펴보았다. 그중 하나는 라이트의 제안을 토대로 하여 기능을 진화의 과정에서 선택된 효과(selected effect)로 보고자 하는 견해이고, 다른 하나는 어떤 체계가 가지는 역량에 대한 분석적 설명에서 사용되는 하부 역량으로 기능을 이해하고자 하는 커밍스의 견해가 그것이다. 이 두 견해는 무엇을 기능으로 간주할 것인가에 대해 상이할 뿐만 아니라 기능적 설명의 대상에 대해 서로 의견을 달리한다.

먼저 무엇을 기능으로 간주할 것인가에 관한 두 견해의 차이에 대해 생각해 보자. 어떤 유기체가 지닌 특질이 한 부류의 환경에서는 자신이 속한 생물종의 유지에 기여하지만 다른 부류의 환경에서는 그렇지 못하거나 오히려 해가 될 수도 있다. 예를 들어, 비둘기의 비행 능력이 변화된 환경에서 어떤 이유로 인해 그 종의 유지에 더 이상 도움이 되지 않거나 오히려 방해가 된다고 하자. 커밍스에 따르면, 이러한 상황에서도 우리는 여전히 비둘기의 비행 능력을 그 날개의 기능으로 간주할 것이다. 그러나, 선택된 효과로서의 기능 개념을 채택할 경우, 변화된 환경에서 비둘기의 비행 능력은 더 이상 그 날개의 기능이 아니라고 말해야 할 것처럼 보인다.<sup>18)</sup> 다른 한편, 언어의 사용은 인간으로 하여금 생존 경쟁에서 매우 유리한 고지를 점유하는데 중요한 기여를 했음이 분명하다. 그런데 이러한 언어사용 행

---

18) 참조: Cummins(1975), 60. 이와 같은 커밍스식의 차별화는, 어떤 특질의 성질이 현시점에서도 그 특질을 가지는 생물종의 선택적 우위를 산출할 때에만 문제의 성질은 기능에 해당한다는 방식으로 선택된 효과로서의 기능 개념을 이해할 때 성립한다. 그러나 선택된 효과로서의 기능 개념이 반드시 그런 방식으로 이해될 필요는 없다.

위는 인간의 언어 획득 능력에 의해 가능해진 것이다. 그러므로 진화의 관점에서 볼 때, 인간의 언어 획득 능력은 선택된 효과에 해당하며, 따라서 라이트 및 그 동조자들은 인간의 언어 획득 능력을 기능으로 간주할 것이다. 그러나 커밍스가 이러한 견해에 동조할지는 분명치 않다. 만약 인간의 언어 획득 능력이 인간의 다른 능력에 대한 분석적 설명에서 하부 능력으로 사용된다면, 전자의 능력은 커밍스의 입장에서 기능으로 간주될 것이다. 이러한 가능성을 인정한다 하더라도 이와 같은 소급이 무한정 반복될 수는 없다. 다시 말해, 어느 단계에서는 그러한 소급이 더 이상 가능하지 않은 인간의 능력들에 부딪히게 될 것이다. 편의상, 인간의 언어 획득 능력이 그러한 능력들 중의 하나라고 하자. 이러한 가정을 채택할 경우, 인간의 언어 획득 능력은 커밍스의 입장에서 볼 때 기능으로 간주될 수 없다. 물론 인간의 언어 획득 능력은 여러 하부 역량들의 복합적 작용에 해당한다. 따라서, 인간의 언어 획득 능력은 관련된 하부 역량들 및 그들 사이의 상호 작용에 의해 분석적 설명이 가능한 대상이 된다. 그리고 이 분석적 설명에서 사용되는 하부 역량들은 커밍스가 기능으로 간주할 대상들이다. 라이트적 관점에서 이 하부 역량들을 기능으로 간주할 여지는 열려 있다. 인간의 언어 획득 능력이 선택된 효과에 해당한다면, 그러한 능력을 산출하는 하부 역량들 역시 선택된 효과로 볼 여지가 있기 때문이다. 물론, 이 하부 역량들은 언어 획득 능력이 선택되는 과정에서 보다 간접적으로 선택된다고 말해야 할 것이다.<sup>19)</sup> 이처럼, 무엇을 기능으로 간주할 것인가라는 물음과 관련하여, 문제의 두 견해로부터 비롯하는 답변은 구체적 적용 사례들을 통해 볼 때 상당히 중복될 여지가 있다. 그러나, 서로 답변을 달리하는

---

19) 실질적인 논의는 이보다 훨씬 복잡하게 이루어질 것이다. 가령, 인간의 언어 획득 능력에 기여하는 어떤 하부 역량(예를 들어, 기억 능력)은 그것이 언어 획득 능력에 기여하는 것보다 훨씬 직접적인 이유로 진화의 과정에서 선택된 경우에 해당한다고 하자. 이런 경우, 문제의 하부 역량이 언어 획득 능력의 산출에 기여함으로써 선택되는 이유는, 그것이 선택되는 직접적인 이유와 비교하여, 파생적이고 부수적인 것에 불과할 수 있다.

사례들도 없지 않을 것이며 이러한 의견 불일치는 두 견해가 개념적으로 상이할 뿐만 아니라 실질적으로 차이가 있음을 시사한다.

나아가서, 3.3과 3.4.1에서 자세하게 논의된 것처럼 이 두 견해는 기능적 설명의 대상에 대해서도 의견을 달리 한다. 즉, 라이트적 견해에서는, 기능을 행하는 특질이나 그러한 특질을 가지는 대상의 존재가 설명의 대상이다. 반면, 커밍스적 견해의 경우, 기능에 해당하는 하부역량들을 토대로 성립하는 어떤 체계의 역량이 설명의 대상이다. 앞에서 이미 언급한 것처럼, 커밍스에 따르면, 기능을 수행하는 특질이나 대상의 존재를 설명의 대상으로 삼는 것은 설명과 최선의 설명으로의 추론을 혼동한데서 비롯하는 그릇된 접근 방식이다.

따라서 이 두 견해는 일견 상호 대립적인 관계에 있는 것처럼 보인다. 즉, 두 견해 중 하나가 옳다면 나머지 견해는 옳을 수 없는 것처럼 보인다. 실제로 두 견해의 주창자 및 지지자들은 자신들의 견해가 그러한 대립적 관계에 있는 것으로 이해하였다. 예를 들어, 커밍스(1975)는 선택된 효과로서의 기능 개념에 대해 매우 적대적이다. 그는 진화의 과정에서 선택된 효과로서의 기능 개념이 생물학에서 기능 언어가 사용되는 많은 경우들을 포획한다는 점을 인정한다. 그러나 이러한 장점에도 불구하고, 커밍스에 따르면, 이 기능 개념은 우리를 “심각하게 오도하며, 생물학 내에서도 그 적용가능성이 극히 제한되어 있다.”(1975, 59) 적용가능성이 극히 제한되어 있다는 비판과 관련하여 커밍스는, 어떤 유기체가 지닌 특질이 한 부류의 환경에서는 자신이 속한 생물종의 유지에 기여하지만 다른 부류의 환경에서는 그렇지 못하거나 오히려 해가 되는 상황을 생각한다. 앞에서 언급한 비둘기의 경우가 그 사례이다. 따라서 기능적 분석은 진화론적 고려와는 별도로 진행될 수 있으며 그렇게 되어야 한다는 것이 커밍스의 생각이다. 뿐만 아니라, 기능 개념의 사용을 정당화하는 것은 어떤 효과들이 설명되는가 하는 것이 아니라 설명의 양식이라는 점에서, 선택된 효과에 초점을 맞추는 접근 방식은 우리를 오도한다. 따라서, 기능에 대한 분석은 진화 과정에서의 선택과는 무관하게 이

루어질 필요가 있으며 그렇게 될 수 있다는 것이 커밍스의 생각이다.

그러나, 기능에 대한 문제의 두 견해가 진정으로 대립적이며 배타적인지는 의심스럽다. 먼저, 어떤 개체가 지닌 역량이 진화의 과정에서 선택된 효과라는 주장과 그러한 역량은 그 개체가 지닌 다른 하부 역량들이 상호 작용한 결과에 해당한다는 주장은 배타적이라기보다는 상호 보완적인 관계에 있는 것처럼 보인다. 커밍스에 따르면, 개체가 어떤 역량을 가진다는 사실은 인과적 관점에서는 “유전적 설계”와 같은 하부 구조에 의해 설명되어야 할 것이다. 그의 주장은 부분적으로는 맞다. 개체가 지닌 특질들은 하부 수준의 물질적 기초에 의존하기 때문이다. 그렇지만, 그 개체의 후손들이 문제의 특질을 계속 지닐 경우, 진화론적 설명이 현재로서는 최선인 것처럼 보인다. 즉, 문제의 특질을 지닌 개체들이 경쟁적 우위를 점하기 때문에, 그 특질은 번식 과정에서 계속 보존된다는 것이다. 이러한 진화론적 설명이 단순히 중복적이거나 불필요한 설명으로 보기 어려운 것은, 개체가 지닌 어떤 특질에 우리가 관심을 가지게 되는 것은 일반적으로 그 개체가 속한 종의 구성원들이 지속적으로 문제의 특질을 가질 때이기 때문이다. 나아가서, 이른바 “근대적 종합” 이후 자연선택에 의한 진화론적 설명과 미시적인 유전인자들에 의한 유전학적 설명은 더 이상 배타적이지 아니라 상호 보완적인 관계에 있는 것으로 이해된다.

실제로 기능에 대한 앞의 두 견해는 배타적이라기보다는 상호 보완적인 성격이 강하다. 먼저, 개체가 지닌 어떤 역량을 진화 과정에서 선택된 효과로 보는 관점은 커밍스 식의 분석적 설명이 지니는 문제점을 보완해 준다. 즉, 커밍스의 분석적 설명은 어떤 특정 시점에서 그 개체가 문제의 역량을 지닌다는 사실을 설명해 주는 역할을 하는 반면, 그 개체의 후손들이 문제의 역량을 계속 지니는 사실을 제대로 설명해 주지 못한다. 라이트 식의 기능적 설명에서 중요한 역할을 하는 것은 바로 후자의 사실에 대한 진화론적 이해이다. 뿐만 아니라, 커밍스의 기능 개념은 다소 방만한 적용을 허용하는 것처럼

보인다. 예를 들어, 인간의 세포 조직은 경우에 따라 종양을 형성하기도 한다. 그 원인을 조사한 결과, DNA 배열에서의 변형 때문인 것으로 밝혀졌다고 하자. 이러한 상황에서 종양의 형성을 변형된 DNA 배열의 “기능”에 의해 설명하는 것은, 키처(Philip Kitcher 1993, 272)의 지적처럼, 기능 개념의 부적절한 적용으로 보인다. 커밍스의 기능 개념이 드러내는 이와 같은 문제점은 “선택된 효과”의 기준을 채택함으로써 해결될 수 있는 것처럼 보인다. 다른 한편, 개체가 지닌 어떤 특질에 대해 라이트 식의 기능적 설명이 시도되는 경우를 생각해 보자. 많은 경우, 그것이 진화의 과정에서 선택된 효과에 해당하는지는 일견 분명치 않을 것이다. 커밍스 식의 분석적 접근은 이 점을 규명하는데 유력한 방안이 될 수 있다. 만약 문제의 특질이 지닌 어떤 역량이 진화의 과정에서 경쟁적 우위를 제공하는 그 개체의 어떤 역량에 기여하는 하부 역량들 중의 하나인 것으로 밝혀진다면, 전자의 특질 역시 보다 간접적인 의미에서 선택된 효과에 해당한다고 할 수 있을 것이다. 왜냐하면 일단의 하부 역량들을 바탕으로 성립하는 어떤 역량이 진화의 과정에서 선택된다 함은 전자의 하부 역량들 역시 선택됨을 함축할 것이기 때문이다.

이 절에서의 논의가 옳다면, 기능에 대한 두 대표적 견해는 양립 가능할 뿐만 아니라 상호 보완적인 관계에 있다고 해야 할 것이다. 그렇다면, 기능이라는 용어가 두 가지 다른 방식으로 규정되고 사용되는 것 역시 이해될 수 있는 것처럼 보인다.

## 5. 기능적 분석의 적용 : 인지과학의 경우

지금까지 다소 일반적 수준에서 이루어진 기능적 분석에 대한 논의가 인지과학의 경우에 어떻게 적용될 수 있는가를 밝히고 그 의의를 평가하는 것이 이 절의 과제이다.

### 5.1 정보 처리 체계로서의 마음

서두에서 언급한 것처럼, 인지과학의 일반적 목표는 마음의 활동, 즉 다양한 정신적 현상들을 이해하는 것이다. 이러한 목표를 추구하는 과정에서, 마음을 일종의 정보 처리 체계로 보는 관점이 폭넓게 채택되어 왔다. 물론 이러한 관점의 채택은 컴퓨터 과학, 특히 인공지능 분야에서의 발달에 의해 커다란 자극과 영향을 받은 것이다.<sup>20)</sup> 이러한 영향은, 길지 않은 인지과학의 역사에서 가장 폭넓게 받아들여온 가정, 즉 “인간의 마음은 정보를 받아들이고 저장하며, 추출하고 변형시키는 동시에 전달하는 복합 체계이다”<sup>21)</sup>는 가정에서 잘 드러난다. 인공 지능의 소유자로서의 컴퓨터는 가장 전형적인 정보 처리 체계이며, 방금 언급된 가정은 다시 말해 인간의 마음 역시 정보 처리 체계라는 가정에 해당한다.

여기서, 컴퓨터가 정보 처리 체계라는 주장의 의미를 좀더 분명히 할 필요가 있다. 컴퓨터의 이론적 원조에 해당하는 튜링 기계는 매우 단순하면서도 강력한 정보 처리 체계이다. 심지어, 연결주의가 부흥하기 이전의 고전적 견해에 따르면, 정보 처리란 튜링 기계들이 행하는 부류의 일들에 해당한다. 따라서, 정보 처리의 의미를 규명하는 한 가지 유력한 방식은 튜링 기계의 구성과 그 작동 방식을 살펴보는 것이다. 먼저, 튜링 기계는 크게 두 부분으로 나뉜다. 하나는 무한히 긴 것으로 가정되는 테이프이고, 다른 하나는 그 테이프에 담겨 있는 내용을 조작하는 기계 헤드(machine head)이다. 전자의 테이프는 다시 일련의 칸들(cells)로 나뉘는데, 각 칸은 많아야 하나의 기호를 담을 수 있다. 기계 헤드 부분은 테이프를 따라 오른쪽 또는 왼쪽으로 한 번에 한 칸씩 움직일 수 있는 구동 장치, 머리 부분 바로 앞에 놓여 있는 칸 속의 기호를 읽을 수 있는 장치, 기계의 현 상태를 등록하는 장치 및 머리 부분이 수행할 작용들을 규정하는 지침들로

20) 참조: Gardner(1985), 119.

21) Stillings, et al.(1995), 1.

구성되는 기계표(machine table)를 포함한다.

이러한 구성을 바탕으로 한 튜링 헤드의 작동 방식은 다음과 같다. 먼저, 프로그래머가 어떤 질문을 튜링 헤드가 알아볼 수 있는 기호들로 테이프 위에 쓴다. 그리고 나서, 특정한 상태에 있는 기계의 헤드를 테이프 위의 특정한 칸 앞에 놓는다. 그러면, 기계는 기계표 속의 지침들을 따라 테이프의 좌우로 움직이는데, 이 과정에서 그것이 읽은 칸의 기호를 다른 기호로 바꾸어 놓기도 한다. 마침내 “정지” 지침에 따라 기계는 작동하기를 멈추는데, 이때 테이프 위에 쓰여 있는 기호들이 주어진 질문에 대한 답을 나타낸다.

그리고 이와 같은 구성과 작동 방식을 통해 튜링 기계는 다양한 질문들에 답할 수 있다. 튜링 기계의 계산 능력을 규정하는데 이론적 근거로 사용되는 처치-튜링 입론에 따르면, “효과적 절차(effective procedure)”라 불려질 수 있는 모든 과정은 어떤 튜링 기계에 의해 구현될 수 있다. 다시 말해, 이는 모든 계산 가능한 것은 보편적 튜링 기계에 의해 계산 가능하다는 주장에 해당한다. 따라서, 고전적 관점에서 볼 때, 보편적 튜링 기계는 정보 처리에 대한 가장 강력한 기술을 제공한다.

## 5.2 세 가지 수준의 가설

앞에서 언급한 것처럼, 인지과학의 목표는 마음의 작동 기제와 원리들을 밝혀내고 이를 통해 다양한 정신 현상을 설명하는 것이다. 이러한 과제의 수행에서, 마음을 정보 처리 체계로 보는 관점이 지니는 함의는 무엇인가? 이 물음에의 답을 위한 단서로서 튜링 기계의 경우를 먼저 생각해 보자.

튜링 기계에 대해서는 하나 이상의 기술이 가능한 것처럼 보인다. 먼저, 그것에 대해 물리적 수준의 기술을 하는 것이 가능하다. 앞에서 이미 논의한 것처럼, 튜링 기계의 행위는 몇 가지 형태의 기본적인 행위들로 이루어진다. 하나는 테이프를 따라 움직이는 것이다. 다른

하나는 테이프 위의 기호들을 읽는 것이다. 또 다른 하나는 테이프 위에 새로운 기호를 기입하는 것이다. 이들 외에도 우리는 튜링 기계의 다른 물리적 성질들을 기술할 수 있다. 예를 들어, 그것의 무게, 테이프의 칸들 사이를 움직이는데 걸리는 시간 등에 대해 기술할 수 있다.<sup>22)</sup> 이와 같은 물리적 기술의 장점은 다른 종류의 기술을 채택했더라면 가능하지 않았을 예측들을 가능하게 한다는 것이다. 그러나, 튜링 기계에 대한 물리적 기술은 그 나름의 원리적 한계를 안고 있다. 첫째, 어떤 튜링 기계에 대한 물리적 기술들 중 극히 일부만이 그것의 정보 처리 기능과 유관하다.<sup>23)</sup> 그런데, 이 소수의 물리적 기술들을 식별해 내는 것은 지난한 일이다.<sup>24)</sup> 둘째, 어떤 튜링 기계가 수행하는 정보 처리 기능들은 물리적으로 상이한 다수의 튜링 기계들에 의해서도 수행될 수 있다. 소위 기능주의자로 불리는 철학자들이 애용해온 복수 실현가능성의 논변이 그것이다.

물리적 기술의 이러한 한계를 극복하기 위해서는, 물리적 장치로서의 튜링 기계가 구현하는 절차상의 원리들에 호소하는 것이 불가피하다.<sup>25)</sup> 튜링 기계의 관점에서 볼 때, 절차상의 원리에 해당하는 것은 우선 그것의 기계표이다. 이미 언급한 것처럼, 튜링 기계의 기계표는 그것이 수행할 정보 처리 작용을 위한 기본적 지침들로 구성되어 있다. 기계표는 튜링 기계의 물리적 구성에 대한 어떤 언급도 포함하지 않지만, 튜링 기계의 상태와 그것이 읽는 기호가 주어지면 그것이 취할 행동을 예측할 수 있게 한다. 그런데 보편적 튜링 기계의 관점에서 보면, 하나의 튜링 기계가 그것의 기계표에 의해 수행하게 될 정보 처리 작용은 그것과 다른 기계표를 가지는 튜링 기계에 의해 수행될 수 있다. 따라서, 보편적 튜링 기계의 기계표를 구성하는 절차상 원리들은 개별 튜링 기계의 기계표를 구성하는 절차상의 원

---

22) 참조: Dawson(1998), 25.

23) 참조: Pylyphyn(1984), 56.

24) 참조: Arbib(1972), 11-12.

25) 참조: Arbib(1972), 11-12 & Dawson(1998), 27.



리들보다 더 일반적(또는 추상적)이다. 실제로 보편적 튜링 기계 속에서는 개별 튜링 기계의 기계표를 구성하는 지침들이 튜링 기계에게 주어지는 질문이나 그것이 도출해 내는 답과 함께 테이프 위에 기입된다.<sup>26)</sup> 반면에, 보편적 튜링 기계의 작동 원리들에 대한 기술은 모든 보편적 튜링 기계들에 공통된 것이다.

그러면, 튜링 기계에 대한 물리적 기술과 절차상의 기술만으로 충분한가? 물리적 기술과 더불어, 절차적 기술 역시 그 나름의 한계를 지닌다. 예를 들어, 두 정수의 합을 계산하는 역할을 수행하는 튜링 기계를 생각해 보자. 이 튜링 기계는 출발점에서 물음에 해당하는 자료가 기재된 테이프와 특정한 지침들로 구성되는 기계표로 구성될 것이다. 그리고 일련의 작업 과정을 거쳐 정지하는 단계에 이르렀을 때, 문제의 튜링 기계는 답에 해당하는 자료가 기재된 테이프와 동일한 지침들로 구성된 기계표로 이루어질 것이다. 그런데, 이 튜링 기계에 대한 절차상의 기술 속에는 그것의 구성이나 작업 과정의 의미에 대한 이야기가 전혀 포함되지 않는다. 이것은 튜링 기계에 의한 정보 처리가 순전히 형식적으로 이루어짐을 말한다. 즉, 테이프 위의 특정한 칸에 쓰여진 기호를 읽고 기계표 속의 지침에 따라 행동을 취할 때 튜링 기계는 그것이 읽는 기호의 의미에 대해 묻지 않는다. 정보 처리의 각 단계에서 튜링 기계의 행동은 그것이 읽게 되는 기호의 모양, 기계의 상태 그리고 연계된 특정한 지침에 의해서 결정된다. 여기서 문제의 튜링 기계에게 주어지는 물음과 그것에 대한 답으로 구성되는 순서쌍들의 집합을 생각해 보자. 주어진 물음들에 대하여 앞의 튜링 기계와 동일한 답들을 산출하는 상이한 튜링 기계들이 존재한다. 물론, 이 튜링 기계들은 그것들의 행동을 결정하는 지침들을 구성하는 기계표가 서로 다르다는 점에서 상이하다. 결과적으로, 개별 튜링 기계에 대한 절차적 기술은 일단의 상이한 튜링 기계들이 주어진 물음들에 대하여 왜 동일한 답들을 산출하는지를 설명하지 못한다. 보편적 튜링 기계의 수준에서도 절차적 기술만으로는 이 문

26) 참조: Dawson(1998), 18 & Minsky(1972), sec. 7. 2.

제가 해결되지 않는다는 것은 비교적 쉽게 추론할 수 있다.

이 문제를 해결하기 위해서는 절차적 기술보다 더 추상적인 기술이 요구된다. 위에서 든 예에서, 동일한 물음들에 대하여 동일한 답들을 산출하는 일단의 튜링 기계들은 두 정수의 합을 구하는 계산을 행한다는 점에서 공통적이다. 문제의 튜링 기계들이 “두 정수의 합을 구하는 계산을 한다”는 기술을 편의상 계산적 기술이라고 부르자.<sup>27)</sup> 계산적 기술은 정보 처리 체계로서의 튜링 기계가 어떤 종류의 문제를 풀고 있는지에 대해 말해 준다. 다시 말해, 튜링 기계에 대한 계산적 기술은 그것이 행하는 계산에 대한 일종의 해석이다. 물리적 기술이나 절차적 기술의 수준에서 상이한 일단의 튜링 기계들이 동일한 물음들에 대해 동일한 답들을 산출하는 것은 모두 동일한 종류의 계산을 행하고 있기 때문이라는 설명을 제공하는 일과 더불어, 계산적 기술은 물리적 기술이나 절차적 기술에 의존하지 않는 예측을 가능하게 한다. 예를 들어, 어떤 튜링 기계가 두 정수의 합을 구하는 계산을 한다는 기술은, 그 기계가 어떤 물질로 어떻게 만들어지며 그것의 기계표가 어떤 지침들로 구성되는가에 대한 정보 없이, 어떤 두 정수의 합이 무엇인가라는 물음에 대해 그 기계가 어떤 답을 산출할 것인가에 대한 예측을 가능케 한다. 그러나, 튜링 기계를 이해하기 위해 물리적 기술이나 절차적 기술만으로 충분하지 않은 것과 마찬가지로, 계산적 기술 역시 그것만으로 충분하지 않다. 가령, 두 정수의 합을 구하는 계산을 하는 튜링 기계들 중 어느 것이 가장 효과적으로 그 작업을 수행하는가 라는 물음에 답하기 위해서는 각 튜링 기계가 문제의 계산을 수행하는 절차에 대한 기술과 이해가 필요하다.

지금까지 우리는 정보 처리 체계로서의 튜링 기계를 설명하거나 이해하기 위해서는 세 가지 다른 수준에서의 기술이 필요하다고 논의해 왔다. 마음 역시 일종의 정보 처리 체계라면, 그것을 설명하기 위해 우리는 다음의 물음들에 답해야 한다. 먼저 마음은 어떤 정보

27) “계산적 기술”이라는 표현의 사용은 Mars(1982)의 선도를 따른 것이다.

처리 문제들을 다루며, 왜 그런 문제들을 해결하려고 하는가? 둘째, 어떤 단계들을 밟아 그 문제들을 해결하는가? 셋째, 그러한 정보 처리가 이루어지는 물리적 바탕은 무엇인가?

정보 처리 체계로서의 마음은 세계에 대한 정보를 받아들이고 저장하며, 필요에 따라 이를 추출하고 변형시킬 뿐만 아니라 이를 바탕으로 행동한다. 우리는 마음을 통해 이루어지는 정보 처리의 형태들과 그 성격을 정확하게 파악하기를 원하며, 왜 이러한 형태의 정보 처리 활동들이 이루어지는지를 이해하기를 원한다. 많은 경우 세계로부터의 정보는 우리의 감각기관을 통해 물리적 신호 또는 자극의 형태로 주어지며, 이는 다시 마음을 통해 다양한 형태의 표상으로 처리된다. 물론 이 중 일부는 정보 처리 체계 속에 저장되고 필요한 경우에 추출되어 사용된다. 그리고 세계로부터의 정보가 표상으로 파악된다고 하는 것은 이미 앞에서 계산적 수준이라 불렀던 수준에서 정보 처리에 대한 기술이 이루어짐을 말한다. 왜 이러한 정보 처리가 일어나는가? 과학적 탐구에서 일어나는 “왜?”라는 물음은 때때로 “어떻게?”라는 물음으로 이해되고 대답되기도 한다. 그러나 여기서 제기되고 있는 “왜?” 질문은 “어떻게?”에 대한 물음과는 분명 구분될 필요가 있다. 물론, 이러이러한 형태의 정보 처리가 어떤 기제를 통해 또는 어떤 절차를 밟아 이루어지는가를 규명하는 것도 주어진 형태의 정보 처리를 이해하기 위해 필요하다. 그럼에도 불구하고, 절차적 규명만으로는 대답되지 않는 질문의 부분이 있다. 여기서의 “왜?” 질문에 제대로 답하기 위해서는, 정보 처리 체계가 속한 환경에 대한 고려가 필요하다. 컴퓨터와 같은 인공적 정보 처리 체계와 관련하여, 주어진 “왜?” 질문에 답하는 것은 거의 사소한 작업이다. 컴퓨터가 이러이러한 정보 처리 활동을 하는 이유는 일차적으로 그것을 설계한 사람의 의도에서 찾을 수 있기 때문이다. 그러나 인간과 같은 생물종과 관련하여, 주어진 “왜?” 물음에 답하는 것은 그렇게 사소한 작업이 아니다. 유력한 추측은, 인간의 경우 환경에의 적응과 진화 과정에서의 선택을 통해 이러저러한 정보 처리 능력을 가지게 되었

다고 말하는 것이다. 그러나 그럴듯한 추측을 하는 것과 그것을 입증하는 것은 전혀 별개의 문제이다. 마음의 다양한 정보 처리 능력이 인간종의 진화 과정에서 선택된 효과에 해당함을 입증할 수 있는가? 이는 많은 경험과학적 탐구를 요하는 작업일 것이다. 이러한 상황에서 가능한 방법론적 선택은, 유력한 추측을 경험적 가설의 형태로 채택하는 것이다. 그리고 이러한 가설의 채택이 과학적 탐구 과정에서 어떤 결과를 낳는가를 살펴보는 것이다. 일부 인지과학자들이 제안하고 채택해 온 “합리성의 원리”는 이러한 접근의 일환으로 볼 수 있다.<sup>28)</sup> 특히, 인간 마음의 정보 처리 능력을 진화 과정상의 선택된 효과로 보는 관점을 취할 경우, 이는 이 글의 전반부에서 논의한 바 있는 기능적 분석의 한 형태를 인지과학적 탐구의 맥락에서 채택하는 것에 해당한다.

인간의 마음이 어떤 종류의 정보 처리 작업들을 수행하며 왜 그러한 능력을 갖추게 되었는가를 이해하는 것만으로 인간의 마음을 충분히 이해했다고 말할 수 없다. 인간의 마음을 이해하기 위해 이론적 모형을 구성하는 방식을 채택한다고 생각해 보자. 물론 이는 컴퓨터와 같은 인공적 정보 처리 체계 및 특정한 프로그램의 개발을 통해 이루어질 수도 있을 것이다. 그리고 이론적 모형에 해당하는 문제의 인공적 정보 처리 체계가 인간의 마음과 마찬가지로 우리가 관심을 가지는 어떤 종류의 정보 처리를 수행한다고 하자. 이 경우, 주어진 이론적 모형이 인간의 마음에 대해 우리가 원하는 이해를 제공해 준다고 볼 수 있는가? 튜링 기계에 대한 논의에서 이미 언급된 것처럼, 동일한 종류의 정보 처리도 여러 가지 다른 방식으로 이루어질 수 있다. 따라서, 어떤 이론적 모형이 인간의 마음에 의해 수행되는 것

---

28) Newell(1982, 102)이 합리성의 원리라 부르는 것에 따르면, “어떤 행위자가 그의 행동들 중 하나가 그의 목표를 달성하게 해줄 것으로 안다면, 그는 그러한 행동을 선택할 것이다.” Anderson(1990, 28)은 이러한 원리를 한층 더 일반화한 “합리성의 일반 원리(General Principle of Rationality)”를 채택한다. 그의 일반화된 합리성 원리에 따르면, “인지적 체계는 항상 유기체의 행위의 적용을 최적화하는 방식으로 작용한다.”

과 같은 종류의 정보 처리를 행한다는 것을 확인하는 것으로부터 인간의 인지 능력에 대한 완전한 이해를 제공해 준다는 결론을 내릴 수 없다. 어떤 종류의 정보 처리가 이루어지는가를 확인하는 것 못지 않게, 그러한 정보 처리가 어떻게 이루어지는가를 이해하는 것 역시 중요하다. 이는 인간의 마음 속에서 일어나는 정보 처리의 기제와 절차를 밝히는 과제에 해당한다. 이러한 유형의 작업과 관련하여 기능주의가 행한 역할을 강조해도 지나치지 않을 것이다. 형이상학적 입론으로서의 기능주의는, 잘 알려진 것처럼, 정신적 상태가 무엇인가라고 묻는 대신에 그것의 역할, 특히 “감각적 자극, 행위적 산출, 그리고 다른 정신적 상태에 대해 그것이 가지는 인과적 관계”<sup>29)</sup>에 의해 이해하고자 하는 입장이다. 예를 들어, 우리가 고통이라고 부르는 정신적 상태들에 공통된 것은 그것들이 공유하는 어떤 물리적 성질이라기보다는 그것들이 공유하는 기능이라고 생각하는 것이 형이상학적 기능주의자들의 입장이다. 이러한 견해의 방법론적 적용에 해당하는 기능주의는 어떤 체계의 작동이나 그 능력을 설명하는 과정에서 그 체계를 구성하는 부분들이 무엇으로 만들어져 있는가라고 묻는 대신에 그 부분들이 하는 역할에 대해 묻는다. 이러한 방법론적 견해가 특히 인지심리학자들에게 매력적인 것으로 받아들여진 것은 당연하다. 그것은 마음의 작동에 대한 탐구가 그것의 물질적 토대에 대한 탐구로부터 독립적으로 이루어지는 것을 허용하는 것처럼 보였기 때문이다. 이 글의 전반부에서 기능적 분석의 두 번째 유형으로 다루었던 커밍스의 기능적 분석은 바로 이러한 방법론적 기능주의의 명시적 정형화에 해당한다고 볼 수 있다.

방법론적 기능주의로서의 기능적 분석 역시 기본적으로 탐구 대상인 체계를 그것의 부분들로 나누어 이해하고자 하는 접근 방식이다. 그러나 여기서 관심의 초점에 놓여 있는 것은 물질적 부분이라기보다는 일차적으로 기능적 부분들이다. 컴퓨터의 프로그래밍은 이러한 접근 방식의 한 정형에 해당한다. 프로그래머의 입장에서 일차적인

---

29) Block(1978), 172.

관심은 컴퓨터가 어떤 특정한 계산 능력을 가지게 하도록 하기 위해 필요한 부분적 기능들과 그 조직화이다. 그러한 기능들이 하드웨어의 수준에서 어떻게 구현될 수 있는가 하는 것은 그에게 이차적인 문제이다. 마음의 작동을 탐구하는 인지과학자들이 속해 있는 상황은 컴퓨터 프로그래머의 그것과는 물론 상당히 다르다. 그들에게 주어진 과제는 마음이 어떤 정보 처리 능력을 가지게끔 만드는 것이 아니다. 우리가 하고 있는 논의의 단계에서 그들의 과제는, 마음이 이러한 정보 처리 능력을 가진 것으로 확인된 상황에서, 어떻게 그러한 정보 처리가 이루어지는가를 설명하는 것이다. 이러한 과제의 해결을 위해 커밍스식의 기능적 분석은 방법론적으로 매우 유용하다. 그것은 마음의 물질적 토대에 대한 염려를 일단 제쳐놓고 순전히 기능적 관점에서 접근하는 것을 허용하기 때문이다. 특히, 이러한 기능적 분석이 우리가 절차적 수준이라고 불렀던 수준에서 이루어지는 것으로 볼 경우, 이는 또 다른 주목할 만한 방법론적 특성을 가진다. 구문론적 작업과 의미론적 작업의 분리가 그것이다. 계산적 수준의 탐구에서 주된 관심사의 일부였던 정보의 표상적 성격은, 절차적 수준에 오면 일단 분리 가능한 과제가 된다. 두 정수를 더하는 계산의 경우를 생각해 보자. 이 정보 처리는 수들에 관한 것이다. 흔한 방법은 수들을 일단 10개의 아라비아 숫자들 및 십진법을 통해 표현하는 것이다. 여기서 십진법은 수들을 아라비아 숫자로 표상하는 특정한 방식을 제공해 준다. 예를 들어, “23”은  $(2*10+3)$ 이라는 수를 나타낸다. 이제 “23”과 “19”를 더하는 계산 문제가 주어지면, 우리는 덧셈의 특정한 연산 규칙들을 사용해 “42”라는 합을 도출해 낸다. 이러한 계산은 거의 기계적으로, 즉 문제의 숫자들이 수를 나타내는 특정한 방식을 의식하지 않고 단순히 연산 규칙들의 적용을 통해 이루어지는 것처럼 보인다. 튜링 기계의 경우는, 두 정수를 더하는 계산 과정의 기계적 성격을 보다 명백하게 현시한다. “0”과 “1”이라는 두 기호만을 사용하는 튜링 기계의 경우를 생각해 보자. 여기서, 문제의 튜링 기계가 수를 표상하는 한 가지 가능한 방식은, 예를 들어, “3”을 “01110”으

로 나타내는 것이다. 그런데 이러한 방식으로 표상된 두 정수를 더하는 과정에서 튜링 기계가 밟는 절차는 순전히 기계적인 것으로 이해될 수 있다. 즉, “0”과 “1” 또는 그 기호들의 어떤 병렬이 어떤 수를 나타내는가에 대한 고려 없이, 튜링 기계의 행동은 특정한 시점에 그것이 읽는 기호의 물리적 모양과 기계표의 지침(즉, 연산 규칙)에 의해 결정된다. 정보 처리 과정에서 구문론적 작업과 의미론적 작업의 이러한 분리가능성은 컴퓨터적 정보 처리 모형을 통해 인간의 마음을 이해하는 접근 방식이 폭넓게 받아들여진 주요 요인으로 이해될 수 있다. 왜냐하면, 그것은 인간의 마음이 행하는 것과 같은 복잡한 활동이 매우 단순한 기계적 활동들을 바탕으로 이루어질 수 있는 가능성을 열어 놓는 것처럼 보였기 때문이다. 물론 이러한 분리가 인간 마음의 정보 처리 과정에서 실질적으로 이루어지는지 그리고, 이루어진다면, 어느 정도 철저하게 이루어지는가 하는 물음은 선형적으로 답해줄 수 없는 것이다. 우리가 취할 수 있는 태도는, 그러한 분리가능성을 토대로 한 접근을 통해 인간의 마음이 행하는 것처럼 보이는 것까지 활동들이 얼마나 잘 설명될 수 있는가를 살펴보는 것이다.

커밍스식 기능적 분석의 요체는 어떤 체계의 복잡한 활동 능력을 어떤 하부 기능들 및 그 조직화에 의해 설명하는 것이다. 이러한 분석은 실질적으로 어떻게 수행가능한가? 요구되는 것은 일차적으로 기능의 분화(functional differentiation)이다. 기능의 분화를 위해 사용되어 온 매우 일반적 방법의 예로는 기능적 분리(functional dissociation)를 들 수 있다. 기능적 분리의 방법은, 어떤 활동의 상이한 측면들이 동일한 독립 변인에 의해 달리 영향을 받는다는 관찰로부터 이 상이한 측면들이 상이한 하부 기능들에 의해 매개된다는 추정으로 나아가는 것이다. 이러한 방법의 사용은 인지심리학에서 진행되어 온 기억의 기체에 대한 탐구를 통해 잘 예시된다. 장기적 기억의 상실에 관해 1960년대 전까지 일반적으로 받아들여진 견해는, 새로운 기억들의 편입이 기억저장소 속에 저장되어 있는 옛 기억들의 지속적 저장을 저해한다는 것이었다. 이와 관련하여, 1950년대 말에 수행된

일단의 실험에 의해, 인간의 기억 활동에서 현저하게 드러나는 두 현상, 즉 수위 효과(primacy effect)와 최신 효과(recency effect)가 실험 속의 어떤 변인에 의해 달리 영향을 받는다는 것으로 밝혀졌다. 예를 들어, 실험대상자들에게 일련의 낱말들을 들려준 다음 자유 회상(free recall)을 하도록 요구하는 실험에서, 자유 회상을 요구하는 시점을 늦추는 조작을 한 결과, 수위 효과는 영향을 받지 않는 반면 최신 효과는 사라지는 것으로 나타났다. 이러한 실험 결과를 바탕으로 심리학자들은 기억 활동이 두 가지 다른 하부 기능, 즉 단기 기억(short-term memory)의 기능과 장기 기억(long-term memory)의 기능의 구분을 필요로 한다는 주장으로 나아갔다.<sup>30)</sup>

그러나, 마음에 대한 커밍스식의 기능적 분석은 해결해야 할 문제들을 안고 있다. 그중 하나는 마음의 활동에 대해 복수의 기능적 분석이 가능할 경우, 후보 기제들 중 어느 것이 실제로 작동하는 기제인가를 가려내는 문제이다. 주어진 현상들을 설명하는 것으로 간주되는 복수의 이론적 모형들이 제시되는 것은 과학적 탐구에서 자주 일어나는 상황이다. 그런데 이 대목에서 우리의 관심을 끄는 것은, 동일한 입력들에 대해 동일한 출력을 산출하는 복수의 기능적 기제들이 존재하는 것처럼 보이는 상황이다. 기능주의의 입장에서는, 동일한 입력에 대한 동일한 출력을 산출하는 기능적 기제들은 동등한 지위를 가지는 것으로 간주될 수도 있다. 그러나 인지과학적 탐구의 목표 중 하나는, 문제의 출력들이 어떤 절차를 거쳐 산출되는가를 밝히는 것이다. 다시 말해, 인지과학자들은 마음과 “강한 의미에서 동등한(strongly equivalent)” 이론적 모형을 구성하고 확인할 수 있기를 원한다. 이와 관련하여 필리신(Zenon Pylyshyn 1981)은 세 가지 유력한 방안을 제시한다. 그중 하나는 그가 “상대적 복잡성(relative complexity)의 증거”라고 부르는 것을 활용하는 것이다. 정보 처리 체계의 관점에서 볼 때, 어떤 문제들은 다른 문제들보다 해결하기 어렵다. 따라서, 주어진 문제들을 해결의 난이도(또는 복잡한 정도)에

30) 참조: Dawson(1998), 110-111.



따라 순서를 매긴다고 하자. 해결의 난이도에 대한 유력한 증거는 문제들을 해결하는데 걸리는 시간일 것이다. 이러한 상황에서 주어진 정보 처리 체계들이 문제들을 해결하는 절차가 상이할 경우, 그것은 해결에 필요한 시간의 차이로 나타날 것이다. 반면, 정보 처리 체계들이 동일한 절차를 밟아 문제들을 해결한다면 소요되는 시간 역시, 다른 변수들의 영향이 같을 경우, 동일할 것이다. 그러므로, “상대적 복잡성의 증거”는 마음에 대한 어떤 이론적 모형의 기능적 기제 및 그 정보 처리 절차가 마음과 동일한지를 결정하는데 유력한 증거의 역할을 할 수 있을 것이다.

나아가서, 지금까지 논의된 커밍스식의 기능적 분석은 반복적으로 적용 가능하다. 인간의 마음이 행하는 것으로 확인된 어떤 정보 처리 활동을 어떤 하부 정보 처리 기능들과 그 조직화를 통해 분석하는 작업이 이루어졌다고 하자. 이 과정에서 등장한 각 하부 정보 처리 기능에 대해 다시 기능적 분석 작업이 진행될 수 있다. 기억에 대한 연구 사례로 다시 돌아가자. 단기 기억(또는 일차적 기억) 기능과 장기 기억(또는 이차적 기억) 기능의 구분은 후에 작업 기억(working memory)과 의미 기억(semantic memory)의 구분으로 발전하는데, 후자의 기능들은 각기 다른 일단의 하부 기능들 및 그 조직화에 의한 것으로 간주된다. 필리신의 제안처럼, 기능적 분석의 반복가능성은 절차상으로 상이한 기능적 기제들을 구분하는데 유용한 방법론적 자원에 해당한다. 필리신이 언급하는 “중간 상태(intermediate states)”의 증거<sup>31)</sup>나 “상대적 복잡성”의 증거를 바탕으로 구분되지 않는 이론적 모형들이 기능적 분석의 어떤 단계에서 성립한다고 하

31) 이것은, 마음과 그 이론적 모형에서 이루어지는 정보 처리의 중간 상태들에 관한 정보들로서, 마음과 강한 의미에서 동등한 이론적 모형을 확인하는 과정에서 사용 가능한 것으로 필리신(1981, 87-88)이 제안하는 또 다른 유형의 증거이다. 이 증거의 활용은 마음과 이론적 모형의 중간상태들이 일치하는지를 확인하는 방식으로 이루어진다. 마음에서 이루어지는 정보처리의 중간 상태들을 확인하는 주된 방법으로는 “프로토콜 분석(protocol analysis)”이 있다.

자. 각 이론적 모형을 구성하는 일부 하부 기능들에 대한 추가 분석은 새로운 “중간 상태”의 증거나 “상대적 복잡성”의 증거를 활용할 수 있는 기회를 제공한다.<sup>32)</sup>

커밍스식의 기능적 분석은 또 다른 내재적인 문제를 안고 있는 것처럼 보인다. 어떤 체계가 지닌 역량을 그것의 하부 역량들에 의해 설명하는 것이 궁극적으로 만족스러운 설명을 제공해줄 수 있는가 하는 물음이 그것이다. 이 물음에 대한 한 가지 가능한 답은 기능적 분석의 과정을 통해 설명력이 생겨난다고 말하는 것이다. 앞에서 이미 언급된 것처럼, 커밍스 자신은 세 가지 기준에 의해 기능적 분석의 설명력을 평가하는 것이 가능하다고 생각한다. 첫째, 기능적 분석의 설명력은 분석하는 역량들이 분석되는 역량보다 단순할수록 크다. 둘째, 분석하는 역량들과 분석되는 역량들의 유형상 차이가 클수록 설명의 의의 역시 크다. 셋째, 주어진 체계를 구성하는 요소들의 조직화가 복잡할수록 관련된 분석적 설명의 의의가 크다. 결과적으로, 현재 논의되고 있는 기능적 분석의 설명력은 분석의 횟수가 증가할수록 커질 가능성이 많다. 왜냐하면, 분석이 거듭될수록 분석하는 역량(또는 기능)들이 점점 단순해질 가능성이 크기 때문이다. 그리고 유형상의 차이 역시 분석이 거듭되는 횟수에 비례하지는 않더라도 커질 가능성이 많다. 따라서, 설명력의 극대화를 지향하는 방법론적 관점에서 볼 때, 문제의 기능적 분석은 허용되는 한, 반복하는 것이 바람직한 것처럼 보인다.

그러나 이러한 방법론적 전망은 여전히 부수적 문제들을 안고 있다. 먼저, 기능적 분석의 방법은 설명력의 산출과 관련하여 적어도 그 프로그램 내부에서는 극복가능하지 않은 한계를 지니고 있는 것처럼 보인다. 왜냐하면, 어떤 체계의 역량을 그 하부 기능들에 의해 설명하고 후자의 하부 기능을 그것의 하부 기능들에 의해 설명하는 과정을 밟는다 하더라도, 최종 단계의 설명에서 사용되는 기능들은

32) 이것은 필리신(1981, 89)이 “요소 분석(componential analysis)”에 의한 증거라 부르는 것을 활용하는 것에 해당한다.

그 자체 여전히 기능들일 뿐만 아니라 설명되지 않은 채로 남아 있게 될 것이기 때문이다. 나아가서 기능적 분석은 일종의 방법론적 딜레마를 야기하는 것처럼 보인다. 즉, 커밍스의 방법론적 전망대로 기능적 분석의 반복이 설명력을 증가시키는 결과를 낳는다 하더라도, 분석적 작업이 무한정 반복될 수는 없다. 다른 한편, 기능적 분석을 어느 단계에서 멈춘다면, 그러한 방법론적 결정이 어떻게 정당화될 수 있는가 라는 의문이 성립한다. 물론, 마음의 활동과 관련하여 원초적 기능들(primitive functions)이 존재한다면, 이는 인지과학자들이 기능적 분석을 멈출 자연스러운 지점을 제공해줄 것이다. 그러나, 이러한 원초적 기능들이 존재하는가 하는 물음은 선형적으로 답해질 수 없다. 이 대목에서 이론적 모형과 실험 결과들 사이의 상호 작용을 중요시하는 답변이 제시될 수 있다. 즉, 관련된 실험 결과들이 기능적 분리의 필요성을 시사하는가에 따라 분석의 반복 여부를 결정하는 한편, 수행된 분석에 의한 설명력이 미미할 경우 분석의 반복을 안내할 실험을 개발할 필요성이 성립하는 것으로 보는 방안이 그것이다. 이제 이러한 방법론적 제안에 따라 기능적 분석을 일정한 횟수 만큼 거듭한 후 등장하는 이론적 모형이 일정한 수준의 설명력을 지닐 뿐만 아니라 더 이상의 분석을 요구하는 실험 결과들도 존재하지 않는다고 하자. 이러한 상황에서 과학자들은 그 시점까지의 분석이 최종적으로 산출하는 하부기능들을 잠정적으로나마 원초적 기능들로 간주하게 될 것이다. 그리고 이러한 분석 및 그 결과는 절차적 수준에서 이루어지는 탐구의 종착역이 될 것이다.

이러한 방법론적 처방은 그 나름의 현실적 가치를 지니고 있다. 그럼에도 불구하고, 더 이상의 분석을 요구하는 실험 결과들이 존재하지 않는다는 부정적 이유의 제시에서 한 걸음 더 나아가 더 이상의 분석이 가능하지 않는다고 말할 긍정적 이유의 제시가 필요하다는 주장 또는 설명되지 않은 기능들에 의한 “설명”을 통해 설명이 성립한다고 보기 어렵거나 그러한 형태의 설명은 근본적으로 제한된 설명력을 지닐 수밖에 없다는 주장 역시 무시하기 어렵다. 그렇다면,

이러한 문제들의 제기에 한층 더 적극적으로 대응할 수 있는 방안은 없는가? 이 대목에서 요구되는 것은 기능에 의한 설명이 아니라 기능에 대한 설명이다. 그리고 앞에서 언급한 바 있는 사례화에 의한 설명이 여기서 해결의 실마리를 제공해 주는 것처럼 보인다. 기능적 분석의 최종 단계에서 등장하여 설명하는 역할을 담당하는 “원초적” 기능들은 더 이상 분석이 가능하지 않은 것으로 간주된다는 점에서 설명되는 않은 채로 남는데, 그렇다고 해서 더 이상의 설명이 필요하지 않은 지위를 가지는 것은 아니다. 커밍스의 지적대로, 체계  $s$  속에서 하부 요소  $x$ 의 기능이  $F$ 를 행하는 것이라면, 그  $x$ 는  $s$  속에서  $F$ 를 행하는 성향을 가져야 하고, 이는  $x$ 의 행동이 어떤 법칙적 규칙성을 드러낸다고 말하는 것에 해당한다. 따라서, 어떤 요소  $x$ 가 기능  $F$ 를 가진다는 주장은 그 자체 설명을 필요로 한다. 그런데 성향으로서의 기능에 대한 설명은 그것이 사례화되는 방식에 대한 설명을 요구한다. 왜냐하면, 어떤 성향이 사례화되는 방식을 이해할 때, 우리는 그것의 규칙적 발현을 이해할 수 있는 입장에 놓이기 때문이다. 그러나 단순하지 않은 성향에 대한 설명은 일단 그것을 다른 성향들로 분석하는 절차를 거칠 필요가 있다. 그것이 사례화되는 방식을 바로 설명하기가 여의치 않기 때문이다. 커밍스식의 기능적 분석이 요청되었던 것은 이러한 이유에서였다. “원초적 기능”도 성향이라는 점에서 설명에 대한 요구로부터 면제되어 있지 않다. 오히려 이들은 사례화에 의한 설명의 유력한 후보이다.

사례화에 의한 설명의 주요한 특성은, 앞에서 이미 언급된 것처럼, 주어진 정보 처리 체계에 대한 기능적 분석의 과정에서 채택된 해석이 더 이상 유지될 수 없는 방식으로 이루어진다는 것이다. 컴퓨터에 대한 기능적 분석의 최종 단계에서 도달하게 될 단순한 기능의 유력한 후보들 중의 하나는 “연접 게이트(and-gate)”이다. 그런데 이 연접 게이트에 대해 가능한 사례화 설명은 그것의 기능을 전기 회로 및 그 구성 요소들에 의해 설명하는 것이다. 그리고 이러한 설명은 연접 게이트에 도달하는 과정에서 사용되어 온 기존의 해석을 포기함으로써

가능해진다. 세 가지 수준을 구분하는 관점에서 볼 때, 정보 처리 체계에 대한 기능적 분석의 최종 단계에서 요청되는 것처럼 보이는 사례화에 의한 설명은 기능적 분석이 이루어지는 절차적 수준의 범위를 넘어서는 것이다. 다시 말해, 정보 처리 체계의 원초적 기능들에 대한 설명은 그 기능이 어떻게 물질적으로 구현되는가에 대한 설명에 해당하며, 이러한 작업은 소위 “실행적 수준(implementational level)”에서 이루어진다.<sup>33)</sup> 이 실행적 수준의 작업은 기능주의의 관점에서는 소홀하게 다루어진 부분이다. 기능적 기제의 물질적 구현은 다양한 방식으로 이루어질 수 있으며, 따라서 기능이 어떤 물질에 의해 어떤 방식으로 구현되는가 하는 문제는 마음의 인지 활동을 이해하고자 하는 과학자의 관심사가 아니라 하는 것이 그 주된 이유였다. 그러나 어떤 체계의 역량들에 대한 설명은 궁극적으로 사례화를 통해 이루어져야 하며 기능적 분석은 사례화에 의한 설명을 위해 필요한 준비 과정이라는 관점에서 본다면, 기능주의자의 시각은 마음에 대한 이해를 불완전한 채로 남겨두는 결과를 낳는다. 여기서 어떤 체계의 완전한 이해를 위해서는 그것의 기능들에 대한 사례화 설명이 실행적 수준에서 이루어질 필요가 있다는 입장은 방법론적 물리주의의 한 적용이라고 볼 수 있다.

## 6. 맺는 말

이 글에서 우리는 기능에 대한 분석의 대표적 유형 두 가지, 즉 라이트식의 분석과 커밍스식의 분석을 소개하고 이들 사이의 관계를 상호보완적인 것으로 볼 것을 제안하였다. 그리고 이러한 방법론적 전망이 인지과학적 탐구에서 실질적으로 어떻게 구현될 수 있으며, 그 의의와 한계는 무엇인가를 밝히고자 하였다.

기능에 대한 라이트식의 분석과 커밍스식의 분석은 흔히 대립적인

---

33) “실행적 수준”이라는 용어의 사용에 대해서는, 참조: Dawson(1998), 7장.

것으로 이해된다. 그러나, 이 두 견해는 기능에 대한 두 가지 다른 질문에 답하는 것으로 볼 필요가 있다. 먼저, 기능을 선택된 효과로 보는 라이트의 견해는 무엇을 기능으로 볼 것이며 왜 그러한 기능이 존재하는가 라는 물음에 답하고자 하는 것으로 이해될 수 있다. 그리고 이 “왜?”라는 물음에 답하려면, 문제의 기능을 가지는 체계 자체만이 아니라 그것이 속한 환경(또는 세계) 및 양자 사이의 관계에 대한 고려가 불가피하다. 반면, 어떤 체계의 역량을 설명하는데 기여하는 하부 역량들을 기능으로 보는 커밍스의 견해는 주어진 체계가 어떻게 그러한 역량을 가지는가 라는 물음에 답하고자 하는 것으로 이해될 수 있다. 그리고 이 “어떻게?”라는 물음에 답하려면, 주어진 체계의 구성 요소들 및 그들 사이의 관계를 고려할 필요가 있다. 결국 기능에 대한 라이트의 견해와 커밍스의 견해는 두 가지 다른 질문, 즉 “왜?”라는 질문과 “어떻게?”라는 질문에 답하고자 하는 시도일 뿐만 아니라 기능을 제대로 이해하기 위해서는 그 두 가지 질문들에 대한 답이 모두 필요하다는 점에서 상호보완적이다.

다음으로, 마음에 대한 인지과학적 탐구에서 기능적 분석이 어떤 역할을 하는가 라는 물음과 관련하여, 우리는 먼저 정보 처리 체계로서의 튜링 기계에 대한 완전한 이해를 위해서는 세 가지 다른 수준(즉, 계산적 수준, 절차적 수준, 그리고 실행적 수준)에서의 작업이 필요하며 각 수준에서 이루어지는 기술(즉, 계산적 기술, 절차적 기술 및 물리적 기술)은 각기 독자적인 역할을 담당한다는 입장을 개진하였다. 이러한 방법론적 관점을 마음에 대한 인지과학적 탐구에 적용하면, 우리는 세 가지 질문에 답할 필요가 있다. 첫째, 마음은 어떤 종류의 정보 처리 능력들을 가지며 왜 그러한 능력들을 가지는가? 둘째, 마음은 어떻게 그러한 정보 처리 능력들을 가지는가? 또는 마음이 행하는 정보 처리는 어떤 기제 및 절차를 통해 이루어지는가? 셋째, 마음에서 이루어지는 갖가지 정보 처리는 어떤 물리적 토대 위에서 이루어지는가? 첫 번째 물음에 대한 답은 우리가 계산적 수준이라고 불렀던 수준에서 이루어지는데 이 과정에서 인지과학자

들은 자주 “합리성의 원리” 또는 그 유사 원리를 채택한다. 그런데, 이러한 접근 방식이 마음의 정보 처리 능력을 진화 과정에서 선택된 효과로 보는 관점과 접목될 경우, 이는 기능에 대한 라이트식의 분석이 인지과학적 탐구의 맥락에서 채용되는 것으로 볼 수 있다. 두 번째 물음에 대한 답은 절차적 수준에서 이루어지는데, 이 과정에서 커밍스식의 기능적 분석이 폭넓게 사용된다. 즉, 마음의 갖가지 정보 처리 역량은 그 하부 역량들 및 그 조직화에 의해 설명된다. 그러나, 절차적 수준에서의 기능적 분석은 복수의 이론적 모형들 중에서 마음과 강한 의미에서 동등한 모형을 가려내는 확인의 문제(problem of identification)와 기능적 분석의 딜레마라고 앞에서 불렀던 문제에 부딪히는데, 이 문제들의 해결을 위해서는 원초적 기능들의 물리적 구현에 대한 실험적 수준에서의 해명, 즉 세 번째 물음에 대한 답이 크게 도움이 되거나 필수적인 것처럼 보인다. 이러한 상황은 기능적 분석이 마음의 인지과학적 탐구에서 중요한 방법론적 역할을 수행함에도 불구하고 그 나름의 한계를 가진다는 사실을 드러낸다. 다시 말해, 방법론적 관점에서 볼 때, 절차적 수준에서 이루어지는 기능에 의한 기능의 분석은 그 자체만으로는 불완전하며, 위로는 계산적 수준에서의 “합리적” 분석(또는 진화론적 해명) 그리고 아래로는 실험적 수준에서 이루어지는 기능의 물리적 토대에 대한 해명을 필요로 한다.

결과적으로, 마음에 대한 인지과학적 탐구에서 기능에 대한 라이트식 견해와 커밍스식 견해는 각각 다른 수준에서 방법론적으로 중요한 역할을 담당할 뿐만 아니라 상호보완적 관계에 있다는 사실을 확인할 수 있다.

## 참고문헌

- Anderson, John R.(1990), *The Adaptive Character of Thought*, Erlbaum.
- Arbib, M.A. (1972), *The Metaphorical Brain*, John Wiley & Sons.
- Block, N. (1978), "Toubles with functionalism". Reprinted in N. Block (ed.), *Readings in the Philosophy of Psychology*, Vol. 1, Harvard University Press, pp.268-305.
- Boorse, C. (1976), "Wright on functions", *Philosophical Review* 85: 70-86.
- Cummins, R. (1975), "Functional Analysis", *Journal of Philosophy* 72: 741-765. Reprinted in E. Sober (ed.)(1994), pp.49-69.
- Cummins, R. (1983), *The Nature of Psychological Explanation*, MIT Press.
- Dawson, M.R.W. (1998), *Understanding Cognitive Science*, Blackwell.
- Gardner, H. (1985), *The Mind's New Science: A History of the Cognitive Revolution*, Basic Books.
- Hardcastle, V.G. (1996), *How to Build a Theory in Cognitive Science*, State University of New York Press, especially, chapter 4.
- Haugeland, J. (1978), "The Nature and Plausibility of Cognitivism", *The Behavioral and Brain Sciences* 1: 215-226. Reprinted in J. Haugeland (ed.)(1981), pp.243- 281.
- Haugeland, J. (ed.)(1981), *Mind Design*, MIT Press.
- Hempel, C. (1959), "The logic of functional analysis", in L.



- Gross (ed.)(1959), *Symposium on Sociological Theory*. Reprinted in C. Hempel (1965), *Aspects of Scientific Explanation*, Free Press.
- Hull, D. & Ruse, M. (eds.)(1998), *The Philosophy of Biology*, Oxford University Press.
- Kitcher, P. (1993), "Function and Design", in P. A. French, et al. (eds.)(1993), *Midwest Studies in Philosophy*, xviii, Univ. of Notre Dame Press. Reprinted in D. Hull & M. Ruse (eds.)(1998), pp.258-279.
- Mars, D. (1982), *Vision*, Freeman.
- Minsky, M.L. (1972), *Computation: Finite and Infinite Machines*, Prentice-Hall.
- Nagel, E. (1961), *The Structure of Science*, Ch. 12, section 1.
- Neander, K. (1991), "Functions as Selected Effects: The Conceptual Analyst's Defence", *Philosophy of Science* 58: 168-84.
- Newell, A. (1982), "The knowledge level", *Artificial Intelligence* 18: 87-127.
- Pylyshyn, Z.W. (1981), "Complexity and the Study of Artificial and Human Intelligence", in J. Haugeland (ed.)(1981), pp.67-94.
- Pylyshyn, Z.W. (1984), *Computation and Cognition*, MIT Press.
- Sober, E. (1993), *Philosophy of Biology*, Westview.
- Sober, E. (ed.)(1994), *Conceptual Issues in Evolutionary Biology*, 2nd ed., MIT Press.
- Stillings, N. A., et al. (1995), *Cognitive Science: An Introduction*, 2nd ed., MIT Press.
- Wright, L. (1973), "Functions", *Philosophical Review* 82: 139-168. Reprinted in E. Sober (ed.)(1994), pp.27-47.
- Wright, L. (1976), *Teleological Explanations*, Berkeley.