

Γ. ΜΑΡΑΓΚΟΣ

ΤΙ ΘΑ ΜΠΟΡΟΥΣΕ ΝΑ ΜΑΘΕΙ Η ΕΜΒΡΙΘΗΣ ΑΥΤΕΓΚΕΦΑΛΟΣΚΟΠΟΣ;

1. Ο μύθος: η φιλοσοφική φαντασία επιφύλασσε στη Μαίρη ασυνήθιστη μοίρα. Από τη γέννηση της ζούσε σε ένα «λευκό» δωμάτιο. Τα μόνα «χρώματα» που μπορούσε να δει ήταν το άσπρο και το μαύρο. Με τον κόσμο επικοινωνούσε με ασπρόμαυρη τηλεόραση. Οι πρακτικές λεπτομέρειες δεν έχουν σημασία. Ενδεχομένως έφερε μονίμως γυαλιά με κρύσταλλα που φιλτράριζαν όλα τα άλλα χρώματα. Απτότητα, αξιοποίησε όλες τις ευκαιρίες, σπούδασε νευροφυσιολογία και κατόρθωσε να γίνει η εξέχουσα αυθεντία στον τομέα της. Η Μαίρη λοιπόν, στο άχρωμο δωμάτιο της, γνώριζε τα πάντα για τις λειτουργίες και για τις καταστάσεις του κεντρικού νευρικού συστήματος. Πράξη δεύτερη: η Μαίρη, ελεύθερη πια βγίνει στον κόσμο (ή αποκτά έγχρωμη τηλεόραση). Ανακαλύπτει τότε πώς είναι να βλέπεις τις τομάτες κόκκινες, τον ουρανό γαλανό, κλπ. Στο σώμα των αρχικών πληροφοριών, όσων διέθετε πριν από την απελευθέρωση της, έρχονται τώρα να προστεθούν και άλλες - το πώς είναι να βλέπεις κόκκινο, κλπ. Επιμύθιο: το αρχικό σώμα γνώσεων της Μαίρης δεν ήταν πλήρες. Η πλήρης, συγχρονική και διαχρονική, νευροφυσιολογική περιγραφή των λειτουργιών και των καταστάσεων του κεντρικού νευρικού συστήματος δεν κάλυπτε το χαρακτηριστικό ποιοτικό περιεχόμενο του συνειδητού αισθητηριακού βιώματος. Φιλοσοφικό συμπέρασμα: ο φυσικοκρατικός υλισμός, δηλαδή η θέση ότι είναι εφικτή η πλήρης νευροφυσιολογική αναγωγή - η πλήρης περιγραφή και εξήγηση - του βιώματος, είναι ψευδής, επειδή ακριβώς αδυνατεί να συλλάβει τις ουσιώδεις, συστατικές, ποιοτικές εκφάνσεις του βιώματος. Ερμηνευτική αιτιολόγηση: η νευροφυσιολογία, όπως εξάλλου και κάθε φυσικοκρατική θεωρία, περιγράφει, εξηγεί και προβλέπει μόνο ο,τι μπορεί να διαδραματίσει αιτιακό μέρος σε κατάλληλες ακολουθίες από συμβάντα· εν προκειμένω, ο,τι μπορεί να είναι αίτιο και/ή αποτέλεσμα διεργασιών στο κεντρικό νευρικό σύστημα. Τα ποιοτικά γνωρίσματα του βιώματος όμως είναι αιτιακά αδρανή - ενώ είναι το αιτιακό αποτέλεσμα νευροφυσιολογικών λειτουργιών.

γιών, τα ίδια δεν δρουν ως αίτια άλλων συμβάντων στο κεντρικό νευρικό σύστημα. Πρόκειται απλώς για επιφαινόμενα των νευροφυσιολογικών διεργασιών. Αφού λοιπόν δεν αφήνουν αιτιακό ίχνος, πώς να προσφέρουν λαβή στα εννοιολογικά και στα κλινικά εργαλεία του νευροφυσιολόγου;

Αυτός ο συνδυασμός «νοητικού πειράματος» και φιλοσοφικού σχολίου είναι ένα μέρος από το λεγόμενο «γνωσιακό επιχείρημα» κατά της φυσικοκρατικής / υλιστικής θεωρίας για τη διάνοια, όπως το διατυπώνει ο Frank Jackson, στο πολυσυζητημένο άρθρο «Epiphenomenal Qualia» (*Philosophical Quarterly*, 32 (1982), 127-136). Το κείμενο αυτό μαζί με το εξίσου πολυσυζητημένο «What is it like to be a bat?» του Thomas Nagel (Nagel, T., *Mortal Questions*, Cambridge U. P., (1979), 165-80), αποτελούν σημαντικές προσθήκες στο οπλοστάσιο του εκσυγχρονισμένου δυϊσμού. Είναι αξιοσημείωτο πως δεν πρόκειται για τον κλασικό δυϊσμό της ουσίας της διάνοιας¹, αφού ούτε ο Jackson ούτε ο Nagel κάνουν λόγο για πνευματική ουσία της διάνοιας. Και για τους δυο τα διανοητικά φαινόμενα εκτυλίσσονται στο κεντρικό νευρικό σύστημα (ΚΝΣ), και οι δύο όμως υποστηρίζουν ότι τα φαινόμενα αυτά έχουν γνωρίσματα πέρα από όσα καλύπτουν επαρκώς οι επιστημονικές, φυσικοκρατικές-υλιστικές θεωρίες για το ΚΝΣ. Αυτές τις, ίσως όχι υπερφυσικές, πάντως «υπερεπιστημονικές» εκφάνσεις των διανοητικών φαινομένων, καλούνται να τις νοηματοδοτήσουν φιλοσοφικές θεωρίες². Ο Jackson, όπως προαναφέραμε, επιλέγει την επιφαινοκρατική θεωρία, μόνο

1. Από τις ελάχιστες πλέον πρόσφατες αναβιώσεις του καρτεσιανού δυϊσμού, αξιοθέατη είναι η θαρραλέα μελέτη του John Foster, *The Immaterial Self: A Defense of the Cartesian Dualist Conception of the Mind*, London / N.Y., (1990).

2. Και άλλοι βεβαίως επικρίνουν το υλιστικό πρόγραμμα στις διάφορες εκδοχές του, χωρίς όμως να θεωρούν αναγκαία τη μια ή την άλλη φιλοσοφική θεωρία για την πραγμάτευση των διανοητικών φαινομένων - για ορισμένους από αυτούς θα αρκούσε μια επιστημονική, έστω και φαινομενολογική, ψυχολογία, ή ψυχο-φυσιολογία. Στη χορεία αυτών θα κατατάσσαμε, ενδεικτικά, τον Ned Block (βλ., π.χ., το κλασικό άρθρο του «Troubles with functionalism» στο Block, N. (ed.), *Readings in Philosophical Psychology*, Harvard U.P. (1980), τ. 1, 268-305)· τον Jerry Fodor (βλ., π.χ., το προγραμματικό κείμενο του «Special Sciences (Or: The Disunity of Science as a Working Hypothesis)», *Synthese*, 28 (1974), 97-115)· τον John Searle, γνωστό για τις επιθέσεις του κατά του αναγωγιστικού υλισμού, της γνωσιοεπιστήμης και της Τεχνητής Νοημοσύνης και για τη βιολογικο-υλιστική αντίληψη του· ο Searle διακηρύσσει μάλιστα ότι θα προτιμούσε να αφαιρεθεί τελείως το πρόβλημα των σχέσεων διάνοιας-σώματος από τη δικαιοδοσία των φιλοσόφων για να αποδοθεί στις βιοεπιστήμες, χωρίς όμως να συμπληρώνει τις προγραμματικές δηλώσεις του με σαφή θετική πρόταση για το πώς ακριβώς θα ήταν μια τέτοια βιολογική θεωρία για τη διάνοια (βλ. π.χ., Searle, J., *The Rediscovery of Mind*, MIT Press (1992)).

όμως, σε ο,τι αφορά τα ποιοτικά γνωρίσματα του βιώματος, και όχι για όλα ενγένει τα διανοητικά φαινόμενα, όπως οι παλαιότεροι οπαδοί της. Φαίνεται έτσι να αναγνωρίζει, ως ένα βαθμό, το κεκτημένο της αιτιακο-λειτουργικής θεωρίας για τη διάνοια. Ο Nagel, από την άλλη, δέχεται τη θεωρία των δύο όψεων, όχι ακριβώς στην παλαιά εκδοχή κατά Wundt, αλλά ως «θεωρία των δύο προοπτικών»: διακρίνει πράγματι την υποκειμενική προοπτική του «βιώματος στο πρώτο πρόσωπο» και την αντικειμενική προοπτική του «βιώματος στο τρίτο πρόσωπο» και θεωρεί ότι η πρώτη δεν ανάγεται στη δεύτερη.

Όπως είναι ίσως ήδη φανερό, το ενδιαφέρον μας εδώ θα στραφεί γύρω από το μέρος του γνωσιακού επιχειρήματος του Jackson το σχετικό με τη Μαίρη· δεν θα ασχοληθούμε με τον άλλο ήρωα του επιχειρήματος, τον χρωματικά οξυδερκέστατο Φρεντ: ο Φρεντ βλέπει δύο χρώματα εκεί όπου οι άλλοι βλέπουμε μόνο ένα. Κατά το γνωσιακό επιχειρήμα, κανενός είδους νευροφυσιολογική πληροφορία για το οπτικό σύστημα του Φρεντ ούτε βεβαίως η παρατήρηση ότι συστηματικά και με συνέπεια ταξινομεί τομάτες σε δύο σωρούς με βάση το χρώμα, ενώ οι υπόλοιποι τις συγκεντρώνουμε σε ένα μόνο, δεν μπορεί να μας διδάξει πώς είναι το ιδιαίτερο χρωματικό βίωμα του Φρεντ.

Είναι θεμιτό, κατά τη γνώμη μου, να εστιάζεται το γνωσιακό επιχείρημα στην περίπτωση της Μαίρης, δεδομένου ότι, ούτως ή άλλως, δεν αλλοιώνεται η ουσία του, ενώ παράλληλα, δικαιώνονται διπλά οι προθέσεις του συγγραφέα. Πράγματι, η περίπτωση του Φρεντ είναι προφανώς έκθετη στην εξής ένσταση: αν τα ποιοτικά γνωρίσματα του βιώματος είναι, όπως ισχυρίζεται ο Jackson, αιτιακά αδρανή επιφαινόμενα, πώς γίνεται να ορίζουν την ταξινομητική συμπεροφορά του Φρεντ-εντέλει, είναι ή δεν είναι επιφαινόμενα; Από την άλλη, καθίσταται, νομίζω, πιο ευδιάκριτη η διαφορά από το συγγενικό επιχείρημα του Thomas Nagel, όπως εξάλλου φαίνεται να επιθυμεί και ο ίδιος ο Jackson (ό.π., σσ. 131-2)¹, δικαιολογημένα ίσως, αφού προκρίνει άλλη φιλοσοφική θεωρία από τον Nagel.

Το γνωσιακό επιχείρημα του Jackson έχει επικριθεί από πολλούς για αλληλεπικαλυπτόμενους αλλά και για σφώς διαφορετικούς λόγους. Ορισμένοι, όπως ο Paul Churchland, ελεγχουν την τυπική ορθότητα του

1. Πρβλ. Biro, J.I., «Consciousness and Obhectivity», στο Davies, M. και Humphries, G. W. (eds.), *Consciousness: Psychological and Philosophical Essays*, Οξφόρδη (1993), 178-96, ιδιαιτέρως σσ. 189-93, όπου καταδεικνύεται η κοινή ουσία του γνωσιακού επιχειρήματος κατά Jackson και του επιχειρήματος περί του βιώματος των νυχτερίδων κατά Nagel, στην περίπτωση του Φρεντ.

άλλοι, όπως ο Terence Horgan ή ο Laurence Nemirow, επισημαίνουν εννοιολογικές ασάφειες· μερικοί, όπως ο David Lewis, μολονότι προτείνουν τρόπους ανασκευής, δέχονται πως, αν το γνωσιακό επιχείρημα ήταν έγκυρο, θα υπονόμειε τη φυσικοκρατική προσέγγιση στο πρόβλημα των σχέσεων διάνοιας-σώματος. Από τη δική του σκοπιά, ο Daniel Dennett επισημαίνει ότι οι απόπειρες να εξηγηθούν οι ποιοτικές εκφάνσεις του βιώματος με τη μια ή την άλλη εκδοχή του υλισμού και, κατά μείζονα λόγο η άποψη ότι οι εκφάνσεις αυτές δεν εξηγούνται υλιστικά, προϋποθέτουν σιωπηρά ότι υπάρχουν ποιοτικά γνωρίσματα· αυτή ακριβώς τη σιωπηρή παραδοχή ο Dennett την εκθέτει, την αναλύει εις τα εξ'ων συνετέθει, και εν πολλοίς, την απορρίπτει.

Στη συνέχεια, θα διεξέλθω επί τροχάδην τα επιχειρήματα αυτά - παραλείποντας άλλα-, για να προτείνω ακολούθως μια πρόεκταση στο Gedankenexperiment του Jackson, που θα μου επιτρέψει να διατυπώσω ορισμένες προσωπικές απορίες σχετικά με τα περιφέρμα *qualia*¹.

2. Τι ακριβώς αγνοούσε η έγκλειστη νευροφυσιολόγος; Κατά τον Jackson, της έλλειπαν ορισμένες πληροφορίες για ορισμένες ιδιότητες του βιώματος, ενώ κατείχε όλες τις σχετικές φυσικές πληροφορίες. Ο T. Horgan² επισημαίνει ότι η έννοια της φυσικής πληροφορίας, όπως εμφανίζεται στο κείμενο του Jackson, χρειάζεται διευκρίνιση, και προτείνει τη διάκριση ανάμεσα σε *ρητά φυσικές πληροφορίες*, όσες ανήκουν ή συνάγονται από θεωρητικά επαρκείς φυσικές περιγραφές και εξηγήσεις των διεργασιών της αισθητηριακής αντίληψης και σε *οντολογικά φυσικές πληροφορίες*, όσες αφορούν φυσικές οντότητες και φυσικές ιδιότητες και σχέσεις (150). Κατά τον Horgan, οι οντολογικά φυσικές πληροφορίες, δεν είναι αναγκαίο να εκφράζονται σε φυσικοκρατική γλώσσα. Με βάση τη διάκριση αυτή, ο Horgan ελέγχει την ορθότητα του γνωσιακού επιχειρήματος, επισημαίνοντας ότι στηρίζεται στο διαφορούμενο όρο «πληροφορία» (151). Προσθέτει άραγε κάτι στην κριτική ανάλυση του επιχειρήματος η διάκριση ανάμεσα σε «ρητά φυσικές» και «οντολογικά φυσικές» πληροφορίες; Νομίζω πως όχι, δεδομένου ότι κινείται παράλληλα προς τη διάκριση φυσικοκρατική/μη φυσικοκρατική - βιωματική πληροφορία: ο πρώτος ό-

1. Ο σχολαστικός όρος «quale» (πληθ. «qualia») αντιστοιχεί στον αριστοτελικό όρο «ποιόν», όπως χρησιμοποιείται, λ.χ., στο *Κατηγορίαί*. Επιλέγω εδώ τις συγκριτικά άκομψες περιφράσεις «ποιοτικά γνωρίσματα», «ποιοτικά χαρακτηριστικά», «ποιοτικές εκφάνσεις», υποθέτοντας ότι έτσι διευκολύνεται η κατανόηση.

2. Horgan, T., «Jackson on Physical Information and Qualia», *Philosophical Quarterly*, 34 (1984), 147-52· οι αριθμοί στο κείμενο παραπέμπουν σε σελίδες αυτής της δημοσίευσης.

ρος ανήκει στο λεξιλόγιο της επιστημολογίας και ο δεύτερος της μεταφυσικής. Για το λόγο αυτό εξάλλου, ο Jackson, στη δευτερολογία του, δεν δυσκολεύεται να αφομειώσει την ένσταση του Horgan¹. Θα ήταν φρονιμότερο να αποφευχθεί η ολίσθηση από το επιστημολογικό σχόλιο στον οντολογικό ισχυρισμό και να τονιστεί πως, ακόμα και αν ορισμένες πληροφορίες για διανοητικά φαινόμενα δεν είναι φυσικοκρατικές, αυτό δεν συνεπάγεται ότι οι οικείοι όροι αναφέρονται σε μη φυσικές οντότητες και διεργασίες. Από τη σκοπιά αυτή, το γνωσιακό επιχείρημα, κατά τη γνώμη μου, καθρεφτίζει την «πανουργία της γραμματικής», που δημιουργεί σε ορισμένους την ψευδαίσθηση ότι ο χαρακτήρας της γλώσσας, ορίζει το χαρακτήρα της πραγματικότητας.

Αν θέλαμε να μείνουμε στο θέμα της αμφισημίας του όρου «πληροφορία», θα μπορούσαμε να διερωτηθούμε αν οι πληροφορίες, όσες κατέχει αρχικά και όσες αποκτά μετά την απελευθέρωση της η ηρωίδα μας, αφορούν γενικούς τύπους (types) ή καθέκαστον δείγματα (tokens) συμβάντων ή οντοτήτων, που δεν ταυτίζονται με τα νευροφυσιολογικά αντίστοιχα. Θα πρέπει ίσως να υποθέσουμε ότι ο Jackson έχει κατά νουν επιμέρους δείγματα - όχι το «κόκκινο εν γένει» αλλά «αυτό το κόκκινο, *hic et nunc*» -, ώστε να διασφαλίζεται η υποκειμενική εξατομίκευση που φαίνεται απρόσιτη στις φυσικοκρατικές θεωρίες. Αν όμως ο λόγος είναι για ιδιότητες - ο,τι επιτρέπει να υπαχθούν τα καθέκαστον σε κλάσεις-, ίσως θα πρέπει να σκεφθεί κανείς ότι ο Jackson εννοεί ότι οι ελλείπουσες πληροφορίες αφορούν τύπους. Οι απορίες του πλατωνικού Σοφιστή δεν είναι μακριά. Ο μύθος της Μαίρης δεν διευκρινίζει το σημείο αυτό, όμως το ζήτημα είναι σημαντικό καθώς αφορά, όχι μόνο τη θεωρητική κάλυψη του βιώματος, αλλά γενικότερα τις δυνατότητες και τον τρόπο γνωστικής σύλληψης του ατομικού. Με άλλα λόγια, το ζήτημα σχετίζεται με τη γενικότερη αντίληψη για τις δυνατότητες και τα όρια των επιστημονικών θεωριών. Αν οι θεωρητικοί όροι αναφέρονται σε καθόλου και στοχεύουν στη διατύπωση νομοειδών σχέσεων, θα ακούσε ίσως η στατιστική συσχέτιση - ταύτιση; - ανάμεσα σε τύπους φαινομένων. Αλλά τότε, στην περίπτωση των διανοητικών φαινομένων, θα χρειαζόταν να λυθεί το πρόβλημα της επιστημολογικής υπόστασης των ψυχοφυσιολογικών νό-

1. Jackson, F., «What Mary didn't Know», *Journal of Philosophy*, 83(1986), 291-5. Η θέση του Horgan τελικά απορρέει από τη γενικότερη αντίληψη του ότι η καταλληλότερη προσέγγιση στο πρόβλημα της διάνοις είναι εκείνη της φαινομενολογικής δημόδους ψυχολογίας, όχι του υλισμού: βλ. Horgan, T. και Woodward, J., «Folk Psychology is Here to Stay», στο Lycan, W.G. (ed.), *Mind and Cognition*, Οξφόρδη (1990), 399-420.

μων. Αν το ζητούμενο είναι η μέγιστη δυνατή εξατομίκευση της αναφοράς των θεωρητικών όρων - και *stricto sensu* αυτό ίσως θα απαιτούσε η πραγμάτευση των ποιοτικών εκφάνσεων των βιωμάτων καθέκαστον, η περιγραφή και η εξήγηση των σχετικών φαινομένων θα ενέπλεκε την αναφορά σε μηχανισμούς: σε αιτιακές ακολουθίες συμβάντων¹. Έχει κανείς την αίσθηση ότι αυτός ο προβληματισμός ανακύπτει σε κινούμενο έδαφος: οι ακραιφνείς οπαδοί της φυσικοκρατικής αντίληψης θα υποστήριζαν ότι οι θεωρητικοί όροι αναφέρονται σε τύπους· ορισμένοι από τους οπαδούς του λειτουργισμού θα πρόκριναν τύπους, άλλοι, δείγματα· μερικοί φιλόσοφοι θα προτιμούσαν τύπους για ορισμένα φαινόμενα και δείγματα για άλλα². Η αναζήτηση ερείσματος για την οικοδόμηση αναλυτικής και ερμηνευτικής σκοπιάς είναι μια άλλη ιστορία και δεν θα μας απασχολήσει εδώ.

Οι L. Nemirow³ και D. Lewis⁴ ελέγχουν το γνωσιακό επιχείρημα, προβάλλοντας τις δύο σημασίες του όρου «γνωρίζω»: «γνωρίζω ότι» και «γνωρίζω πώς». Η πρώτη σημασία αναφέρεται στο πληροφοριακό περιεχόμενο του ενεργήματος που δηλώνει το ρήμα ενώ η δεύτερη, παραπέμπει σε ικανότητες του γνωστικού υποκειμένου. Όσες και όσοι θα περίμεναν ότι με την ανάδειξη της δεύτερης σημασίας, θα γινόταν λόγος για πρακτικές ικανότητες της Μαίρης, θα απογοητευτούν - η εμβριθής νευροφυσιολόγος παραμένει, και στην προοπτική των αναλύσεων αυτών απλή θεωρός: της αναγνωρίζεται απλώς η ικανότητα να θυμάται, να λησμονεί, να φαντάζεται, να αναγνωρίζει (Nemirow, ό.π., 491· Lewis, ό.π., 514-8). Όσο για τις ικανότητες να μαθαίνει, και να ανακαλύπτει εμπράκτως φαίνεται πως τις καλύπτει επαρκώς η ικανότητα να διαβάξει συγγράμματα νευροφυσιολογίας και να παρακολουθεί τηλεοπτικά προγράμματα στην ασπρόμαυρη συσκευή της. Ο εγκλεισμός δεν της στερεί μόνο το χρωματικό βίωμα, μοιάζει να αναστέλει και τις τελεστικές λειτουργίες του αισθησιοκινητικού φλοιού του εγκεφάλου της.

Η κριτική από τον P. Churchland, ήταν αφορμή για τη δευτερολογία του Jaeskon (βλ. σελ. 103, υποσημ. 1)· ο Churchland ανταπάντησε ανασυγκρότώντας ακόμη συστηματικότερα το γνωσιακό επιχείρημα

1. Για την ιδέα της εξήγησης μέσω μηχανισμών, βλ. Van Parijs, P., *Evolutionary Explanation in the Social Sciences*, N.J., (1981).

2. Πρβλ., π.χ., την εισαγωγή στο Dennett, D.C., *Brainstorms*, MIT Press (1981) ιδιαίτερος xvi-xx.

3. Nemirow, L., «Physicalism and the Cognitive Role of Acquaintance», στο Lycan, W.G. (ed.), *Mind and Cognition*, Οξφόρδη (1990), 490-9.

4. Lewis, D., «What Experience Teaches», στο Lycan, W.G. (ed.), *Mind and Cognition*, Οξφόρδη (1990), 499-519.

μα¹. Η στρατηγική του έχει δύο αιχμές: αφενός, την αντιστροφή του λεγόμενου (νόμου του Leibniz) για την ταυτότητα των μη διακρίσιμων - ενπροκειμένου «τη μη ταυτότητα των διακρίσιμων»· αφετέρου, την αμφισημία του ρήματος «γνωρίζω». Πιο συγκεκριμένα, στη γραμμή του Churchland, το γνωσιακό επιχείρημα θα μπορούσε να ανασυγκροτηθεί ως εξής (Πρβλ. Churchland [1], 61-3 και Churchland [2], 67-8):

(Π1) Η Μαίρη (πριν από την απελευθέρωση της) γνωρίζει όλες τις φυσικές πληροφορίες για τις καταστάσεις του εγκεφάλου και για τις ιδιότητες τους.

(Π2) Η Μαίρη (πριν από την απελευθέρωση της) δεν γνωρίζει τα πάντα για τα αισθήματα και για τις ιδιότητες τους (αφού μαθαίνει γι αυτά κάτι, μετά την απελευθέρωση της).

(Σ) Τα αισθήματα και οι ιδιότητες τους \neq τις καταστάσεις του εγκεφάλου και τις ιδιότητες τους.

Στη συμβολική γλώσσα του λογισμού κατηγορημάτων με ταυτότητα, ο τύπος:

$$(I) \quad (x)(y)(P)[(Px \leftrightarrow Py) \Rightarrow (x=y)],$$

εκφράζει το νόμο του Leibniz περί ταυτότητας, και θα διαβαζόταν, «για οιοδήποτε x και για οιοδήποτε y , το x ταυτίζεται με το y , αν οτιδήποτε κατηγορείται για το x , κατηγορείται και για το y ». Συνεπώς, αν υπάρχει ιδιότητα τέτοια ώστε να κατηγορείται για το x , χωρίς όμως να κατηγορείται και για το y , το x δεν ταυτίζεται με το y .

Το γνωσιακό επιχείρημα ως ειδική περίπτωση εφαρμογής του νόμου του Leibniz δηλώνει ότι για κάθε διανοητική κατάσταση x και για κάθε εγκεφαλική κατάσταση y , υπάρχει η ιδιότητα «η Μαίρη γνωρίζει όλες τις φυσικές πληροφορίες» και είναι τέτοια ώστε να κατηγορείται για τα y αλλά όχι για τα x . Συνεπώς, $x \neq y$: οι διανοητικές καταστάσεις δεν ταυτίζονται με τις εγκεφαλικές. Συμβολικά, στη γλώσσα του λογισμού των κατηγορημάτων δεύτερης εδώ τάξης (εφόσον στην εμβέλεια καθολικού ποσοδείκτη βρίσκεται μεταβλητή-κατηγορημα):

$$(Γ1) \quad (x)(f)[(Mx \ \& \ Bx) \Rightarrow K(f)mx]$$

$$(Γ) \quad (Γ2) \quad (\exists x)(\exists f)[Mx \ \& \ \sim K(f)mx]$$

$$(\Sigma_{\Gamma}) \quad (\exists x)[Mx \ \& \ \sim Bx]$$

1. Βλ. Churchland, P.M. [1], «Reduction, Qualia, and the Direct Introspection of Brain States», κεφ. 3 στο Churchland, P.M., *A Neurocomputational Perspective. The Nature of Mind and the Structure of Science*, MIT Press (1989), 47-66, ειδικώς το τμήμα «Jackson's Knowledge Argument», 61-6· τ.ί. [2], «Knowing Qualia: A Reply to Jackson», ό.π., κεφ. 4, 67-76. Πρβλ., τ.ί., *Matter and Consciousness*, MIT Press, (1988), 31-4.

όπου «m» = «Μαίρη», «K(f)mx» = «η Μαίρη γνωρίζει ότι το x έχει τη ιδιότητα f»: «Mx» = «το x αφορά κάτι διανοητικό», «Bx» = «το x αφορά κάτι εγκεφαλικό/φυσικό»¹.

Στην εκδοχή αυτή, το γνωσιακό επιχείρημα εμπίπτει στα παράδοξα που ανακύπτουν από μίξη τη του εκτασιακού νόμου του Leibniz με εντασιακά κατηγορήματα, όπως τα «γνώσεως σημαντικά». Είναι γνωστό ότι όσα συμφραζόμενα έχουν εντασιακό κύριο τελεστή - ενπροκειμένω το ρήμα της κύριας πρότασης - είναι «αναφορικός αδιαφανή» (referentially opaque): δεν δηλώνουν μονοσήμαντα όποιο αντικείμενο αφορούν. Ας ανακαλέσουμε στη μνήμη το «παράδοξο της Ηλέκτρας»:

(H1) Η Ηλέκτρα γνωρίζει ότι ο Ορέστης είναι αδελφός της.

(H2) Η Ηλέκτρα δεν γνωρίζει ότι ο άνθρωπος που στέκεται εμπρός της είναι ο Ορέστης.

(H3) Ο άνθρωπος που στέκεται εμπρός της είναι ο Ορέστης.

(Σ_H) Η Ηλέκτρα γνωρίζει και δεν γνωρίζει ότι ο ίδιος άνθρωπος είναι αδελφός της.

Προφανώς, αν και παράδοξο, το συμπέρασμα (Σ_H) είναι «υγιέστερο» από το συμπερασμα (Σ): σχολιάζει τη γνωσιολογική κατάσταση της Ηλέκτρας, δεν εκφράζει οντολογικές δεσμεύσεις. Από το ότι η Ηλέκτρα δεν γνωρίζει ότι ο X είναι ο O, δεν συνάγεται ότι ο X δεν είναι ο O.

Κατ' αναλογία, και αυτό το σκεπτικό επιτρέπει να διακρίνουμε πως το ότι η Μαίρη αγνοούσε ορισμένες διανοητικές πληροφορίες για το βίωμα - εν προκειμένω, τις ποιοτικές-, όσο γνώριζε μόνο τις φυσικές-νευροφυσιολογικές, δεν συνεπάγεται ότι οι διανοητικές πληροφορίες δεν είναι νευροφυσιολογικές. Θα μπορούσε βεβαίως κανείς να αναρωτηθεί περαιτέρω, για το χαρακτήρα, φυσικό ή άλλο, της γνώσης σχετικά με τα διανοητικά φαινόμενα. Η απάντηση δεν μονοσήμαντη. Εξαρτάται από το πώς εκλαμβάνεται το ερώτημα, ως *quid facti?* ή ως *quid juris?* Στην πρώτη περίπτωση, το κεκτημένο των ψυχολογικών και των νευροφυσιολογικών κλάδων δεν επιτρέπει επί του παρόντος τη σύνθεση σε ομογενές εννοιολογικό και μεθοδολογικό συνεχές. Υπό το πρίσμα αυτό, θα ήταν τετριμμένη αλήθεια ότι, στο παρόν στάδιο της γνώσης, «τα διανοητικά φαινόμενα δεν είναι νευροφυσιολογικά». Όποιος όμως θα θεωρούσε το γεγονός αυτό, και την αδυναμία ή την απροθυμία το να φανταστεί πώς θα μπορούσε να είναι διαφορετική η γνωστική κατάσταση, βαθειά ενόραση στην αναγκαία φύση των διανοητικών πραγμάτων, θα διέτρεχε τον κίνδυνο να διαγνωσθεί ως πάσχων από ό,τι ο Daniel Dennett ονομάζει «σύνδρομο των φιλοσόφων»². Και βεβαίως, όσοι, όπως ο Jackson και

1. Πρβλ. Churchland [2], 72-3.

2. Dennett, D.C., *Consciousness Explained*, Λονδίνο/N.Y., (1992), 401.

οι άλλοι φίλοι των ποιότητων, υπεραμύνονται της φιλοσοφίας έναντι του επιστημονικού ιμπεριαλισμού, δεν θα επιχειρηματολογούσαν με μόνο έρεισμα τετριμμένες εμπειρικο-πρακτικές αλήθειες. Η κίνηση από το πεδίο των αληθειών αυτών στη σφαίρα των *quid juris?* ζητημάτων, θα ήταν παράλληλη προς την τάξη των μεταβλητών μέσα την εμβέλεια του καθολικού ποσοδείκτη στη συμβολική αναπαράσταση (Γ) του γνωσιακού επιχειρήματος, στην προηγούμενη σελίδα: από τις ατομικές μεταβλητές - «για κάθε φαινόμενο x» -, έως τις μεταβλητές-κατηγορήματα - «για κάθε ιδιότητα f». Το επόμενο βήμα θα ήταν να γενικευτεί το επιχειρήμα (Γ) ως προς τον τελεστή «Κ». Αξίζει να τονιστεί ότι, θεωρώντας και το «Κ» μεταβλητή, θα βεβαιώναμε κάτι περισσότερο από το «η Μαίρη γνωρίζει τα πάντα σχετικά με...». Θα βεβαιώναμε ακριβώς ότι «η Μαίρη κατέχει κάθε είδους γνώση σχετικά με...»: αντί για την έκταση του ρήματος «γνωρίζω», θα εστιάζαμε στην ένταση του.

Είδαμε πρωτύτερα ότι οι Horgan, Nemirow και Lewis ελέγχουν την ορθότητα του γνωσιακού επιχειρήματος, επικαλούμενοι ακριβώς πολλαπλές και αποκλίνουσες συνδηλώσεις του ρήματος «γνωρίζω» ή των ουσιαστικών «γνώση» και «πληροφορία». Με τη σειρά του ο Churchland φέρνει στο προσκήνιο τις δύο, ήδη από κείμενα του Russell γνωστές, σημασίες «γνώση με περιγραφή» (knowledge by description) και «γνώση με (άμεση βιωματική) γνωριμία» (knowledge by acquaintance). Το πρώτο είδος αντιστοιχεί στη γλωσσικά κωδικευμένη και το δεύτερο στην αισθητηριακά προσλαμβανόμενη γνώση. Με αυτή τη διάκριση στο νου, μπορούμε να πούμε ότι το γνωσιακό επιχειρήμα βεβαιώνει πως η γνώση η κωδικευμένη στη γλώσσα της νευροφυσιολογίας δεν απεικονίζει ούτε είναι δυνατό να απεικονίσει πλήρως τη γνώση που προσλαμβάνεται με άμεση βιωματική εμπειρία. Το πρόβλημα λοιπόν της Μαίρης είναι κατ' ουσίαν το πρόβλημα των σχέσεων ανάμεσα σε δύο τρόπους παράστασης του κόσμου: πώς γίνεται μέσα στον οργανισμό - στη διάνοια; στο κεντρικό νευρικό σύστημα; - η συναρμογή της συμβολικής και της αισθητηριακής παράστασης. Η συναρμογή αυτή έχει άραγε το χαρακτήρα σημασιολογικής σχέσης συμβόλου-περιεχομένου; Συνδέονται άραγε οι δύο παραστάσεις μέσω της αιτιατικής ιστορίας τους ή μήπως πρόκειται για σχετικά αυτοτελείς συνιστώσες του φαινομένου βιώματος με ανεξάρτητες αιτιακές ιστορίες; Ή πάλι, για να χρησιμοποιήσουμε μια τοπολογική μεταφορά, αν δεχτούμε ότι υπάρχει συστηματική απεικόνιση της αισθητηριακής παράστασης στη συμβολική, μπορούμε να αναρωτηθούμε, αν η απεικόνιση αυτή είναι «συνεκτική» - αν, λ.χ., υπάρχουν στην απεικόνιση αυτή κενά, περιοχές άρρητου αισθητηριακού βιώματος, αν έχει το χαρακτήρα «ισομορφισμού» - αν δηλαδή είναι τέτοια ώστε να

διατηρούνται οι δομικές ανάμεσα στα στοιχεία τους. Και βεβαίως θα χρειαζόταν να ασχοληθούμε με την *εκτέλεση* (implementation) των παραστατικών - γλωσσικών/συμβολικών και αισθητηριακών - λειτουργιών του φαινομένου βιώματος στο επίπεδο του οργανικού υποστρώματος. Οι απαντήσεις σε τέτοιου είδους ερωτήματα θα συσχετίζονται βεβαίως με φιλοσοφικές θέσεις και αναλύσεις αλλά θα εξαρτώνται και από τα πορίσματα εμπειρικής μελέτης σχετικά με τις λεπτομέρειες των οικείων διανοητικών λειτουργιών. Ενδεχομένως η πραγμάτευση των σχετικών προβλημάτων απαιτεί και τη συγκρότηση νέου εννοιολογικού πλαισίου, πέρα ό,τι προσφέρουν σήμερα η ψυχολογία, η ψυχοφυσιολογία ή η νευροφυσιολογία. Αυτός είναι για παράδειγμα ο προσανατολισμός στοχαστών όπως ο Paul Churchland¹.

Δεν είναι εδώ η θέση για να αποτιμήσουμε τις προοπτικές αυτές - για τις ανάγκες του προβληματισμού μας θα κρατήσουμε το εξής διδάγμα από τις προηγούμενες παρατηρήσεις: η πολυσημία των γνωστικών όρων δείχνει ίσως ότι κατανοούμε ατελώς τι δηλώνουν. Γενικεύοντας, θα μπορούσαμε ίσως τα υποθέσουμε ότι τη φαινομενολογική γλώσσα, όπου ανήκουν οι γνωσιακοί όροι, τη χαρακτηρίζει ουσιαδώς σημασιολογική απροσδιοριστία - υπόθεση όχι ξένη προς τις γνωστές θέσεις των Quine και Davidson. Κατά ταύτα, το γνωσιακό επιχείρημα δεν είναι αρκετά ειδικό: θα γενικεύοταν για κάθε γλώσσα, όχι μόνο για τη φυσικοκρατική, αλλά και για τη γλώσσα του οντολογικού δυϊσμού και για την καθημερινή, ομιλούμενη γλώσσα. Πράγματι, η ηρωίδα του γνωσιακού επιχειρήματος θα μπορούσε να είναι εμβριθής «εκτοπλασμολόγος» - θα ήταν τότε ικανή να περιγράφει και να εξηγεί όλα τα διανοητικά φαινόμενα αποδίδοντας τα με κάθε λεπτομέρεια στην ανθρώπινη «εκτοπλασματική ουσία». Όμως όταν, για πρώτη φορά, θα αντίκρυζε ώριμες τομάτες, θα μάθαινε κάτι που μόνη η «εκτοπλασμολογία» δεν της είχε διδάξει².

Ο Daniel Dennett, από την πλευρά του, διερωτάται πόσο πολλά γνωρίζει η Μαίρη, όταν γνωρίζει «πάντα» τα νευροφυσιολογικά για τη χρωματική όραση³. Ίσως τα «πάντα» εδώ να σημαίνει πολύ περισσότε-

1. Βλ. λ.χ., Churchland, P. M., *A Neurocomputational Perspective*, κεφ. 5: «Some Reductive Strategies in Cognitive Neurobiology», 77-110· κεφ. 9: «On the Nature of Theories: A Neurocomputational Perspective», 153-96. Πρβλ. Stich., S., *From Folk Psychology to Cognitive Science. The Case Against Belief*, MIT (1983), όπου γίνεται συστηματική κριτική αποτίμηση της φαινομενολογικής δημόδους ψυχολογίας αναφορικά με την έννοια της πεποίθησης και διερευνώνται οι προοπτικές για υπέρβαση του αντίστοιχου εννοιολογικού πλαισίου μέσω ενός γνωσιοεπιστημονικού υποδείγματος.

2. Churchland, P. M., [2], 71.

3. Dennett, D.C., *Consciousness Explained*, Λονδίνο/N.Y., (1991), 398-411.

ρα απ' όσα, διαισθητικά, νομίζουν οι φίλοι των ποιητών¹. Ίσως η Μαίρη να γνωρίζει επακριβώς πώς κάθε χρωματικό ερέθισμα επιδρά στο νευρικό σύστημα της και ποια είναι τα αντίστοιχα νευροφυσιολογικά αποτελέσματα. Θα μπορούσε ίσως να μάθει, «εκ των έndon», ότι κάποιο χρώμα δεν είναι κίτρινο ή δεν είναι κόκκινο, αν είχε επισημάνει τις απολύτως ειδικές αντιδράσεις του εγκεφάλου της, όσες προκαλούνται μόνο από τα κίτρινα ή μόνο από τα κόκκινα ερεθίσματα². Οι οπαδοί του γνωσιακού επιχειρήματος θα μπορούσαν να αντιτάξουν ότι το θέμα δεν είναι αν η Μαίρη πριν την απελευθέρωση της μπορούσε να πει, λ.χ., «όταν κάποτε διεγερθεί κατάλληλα η περιοχή V4 στον κύριο οπτικό φλοιό του εγκεφάλου μου, θα δω κόκκινο», ή κάτι ανάλογο στο πληρέστατο νευροφυσιολογικό λεξιλόγιο της, αλλά αν έχει πραγματωθεί η πλήρης αιτιακή διεργασία με αφετηρία τη διέγερση των χρωματικών υποδοχέων στον αμφιβληστροειδή από φωτόνια κατάλληλης συχνότητας και κατάληξη όποια διέγερση της περιοχής του εγκεφάλου βιώνεται ως χρωματικό αίσθημα. Ακόμη και από αστηρά φυσιοκρατική σκοπιά, το ζήτημα είναι πώς η Μαίρη θα γνώριζε «εκ των έndon» τα αποτελέσματα της διέγερσης του εγκεφάλου της, όσο θα αποκλείοταν συστηματικά η δυνατότητα να πραγματωθούν τα αίτια τους, εν προκειμένω τα κατάλληλα χρωματικά ερεθίσματα;

Στη συνέχεια θα διερευνήσω το ενδεχόμενο να μπορεί η έγκλειστη νευροφυσιολόγος να αποκτήσει, ήδη πριν την απελευθέρωση της, δηλαδή πριν ακόμη διεγερθούν για πρώτη φορά οι περιφερειακοί χρωμοευαίσθητοι υποδοχείς του αμφιβληστροειδούς, χρωματικά βιώματα έτσι ώστε να αυτοανψωθεί ηρωικά έως το επίπεδο της πλήρους χρωματικής γνώσης.

3. Πριν προχωρήσω εμπλουτίζοντας την ιστορία της Μαίρης, θα ήθελα να προτάξω ένα σύντομο σχόλιο σχετικό με τα νοητικά πειράματα. Τις τελευταίες δεκαετίες, τέτοιου είδους νοητικές κατασκευές χρησιμοποιούνται συχνά στη φιλοσοφία της διάνοιας: πλήθος επιχειρήματα αφορούν

1. Θα άξιζε να επισημανθεί εδώ ότι, αν «γνωρίζω τα πάντα» σημαίνει, ευλόγως, «διαθέτω πλήρη θεωρία», πέρα από το πρακτικό ζήτημα του πώς θα συγκροτηθεί και πώς θα είναι η μελλοντική πλήρης νευροφυσιολογική θεωρία, ανακύπτει και το εξής λογικό πρόβλημα: όπως έχει αποδείξει ο Gödel, όσες θεωρίες είναι αρκετά πλούσιες ώστε να εκφράζουν την αριθμητική, και δεν περιέχουν αντιφάσεις, δεν είναι πλήρεις. Αυτή η δυσκολία παρακάμπτεται, αν τεθεί ως αίτημα ότι το σημασιολογικό πεδίο αναφοράς της νευροφυσιολογικής θεωρίας είναι αδύστως ευρύ αλλά πάντως πεπερασμένο. Αυτό όμως το σημασιολογικό αίτημα δεν θα περιόριζε ίσως την εμπειρική επάρκεια της θεωρίας, αν το πεδίο αναφοράς της το συγκροτούσαν γενικοί τύποι φαινομένων· θα ήταν όμως εμπειρικά λιγότερο εύλογο, αν η θεωρία θα στόχευε να καλύψει το απεριόριστο πλήθος των επιμέρους δειγμάτων.

2. Dennett, D.C., ό.π., 400-1.

τον εξωσωματικό βίο εγκεφάλων *in vitro*¹, τις ψυχονητικές λειτουργίες των κατοίκων της «δίδυμης γης»², τα βιώματα ποικίλων ανδροειδών και άλλων τεχνημάτων. Τέτοιες «ιστορίες φιλοσοφικής φαντασίας» δεν αποτελούν πρόσφατη απομίμηση του επιστημονικού συρμού ή ρητορικό εργαλείο για όσους σαγηνεύει το επιστημονικοφανές ύφος. Η αφετηρία αρκετών από τα σύγχρονα επιχειρήματα αυτού του είδους φαίνεται μάλλον να ανήκει σε κλασικό φιλοσοφικό τόπο. Ας θυμηθούμε τη συναρπαστική παράγραφο του Descartes:

[Α]ν υπήρχαν... τέτοιες μηχανές, με τα όργανα και την εξωτερική μορφή πιθήκων, ή άλλων αλόγου ζώων, δεν θα είχαμε κανένα μέσο για να αναγνωρίσουμε πως δεν θα είχαν σε όλα τους την ίδια φύση με τα ζώα αυτά. Ενώ, αν υπήρχαν μηχανές που να μοιάζουν με τα κορμιά μας και να μιμούνται τις πράξεις μας, ..., θα είχαμε πάντα δύο μέσα πολύ σίγουρα για ν'αναγνωρίσουμε πως αυτές οι μηχανές δεν θα ήταν για τούτο αληθινοί άνθρωποι. Το πρώτο μέσο είναι πως δεν θα μπορούσαν ποτέ να μεταχειριστούν την ομιλία, ούτε και άλλα σημάδια, συνθέτοντας τα, όπως κάνουμε εμείς για να εκδηλώσουμε στους άλλους τις σκέψεις μας. Γιατί, μπορεί βέβαια να φανταστεί κανένας μια μηχανή φτιαγμένη έτσι που να προσφέρει λόγια σχετικά με σωματικές ενέργειες που προκαλούν κάποια μεταβολή στα όργανα της, όπως, λ.χ., αν την αγγίζει κανένας σ' ένα σημείο να ρωτά τί θέλουν να της πουν, αν σ' ένα άλλο να φωνάζει πως την πονούν, κι άλλα παρόμοια. Όχι όμως και να τακτοποιεί τα λόγια της με τρόπους διάφορους ώστε ν' απαντά στην έννοια των όσων θα ειπωθούν μπροστά της, όπως μπορούν να κάνουν οι άνθρωποι οι πιο αποβλακωμένοι. Και το δεύτερο μέσο είναι πως, κι αν ακόμα έκαναν πολλά πράγματα όσο καλά, ίσως μάλιστα και καλύτερα από οποιονδήποτε μας, θα νστερούσαν αναπόφευκτα σε μερικά άλλα, από τα οποία θα ανακάλυπτε κανένας πως δεν θα ενεργούσαν ενσυνείδητα, αλλά μονάχα χάρη στη διάταξη των οργάνων τους. Γιατί, ενώ το λογικό είναι ένα όργανο καθολικό, που μπορεί να χρησιμέψει σε κάθε είδους συντυχίες, τα όργανα αυτά [τής μηχανής] έχουν ανάγκη από κάποια ιδιαίτερη διάταξη για κάθε ιδιαίτερη πράξη. (Descartes, R., *Λόγος περί της Μεθόδου*, μτφ. Χ. Χρηστίδη, Αθήνα (1976), 51-2).

1. Βλ., λ.χ., Smart, J. J. C., «Materialism», στο Feigl, H., Sellars, W., Lehrer, K. (eds), *New Readings in Philosophical Analysis*, N. Y. (1972), 378-85. Dennett, D.C., «Where Am I?», στο Hofstadter, D.R., Dennett, D.C. (eds), *The Mind's I*, 219-29 - ο τόμος αυτός είναι εξάλλου πλούσιο αντιπροσωπευτικό ανθολόγιο κειμένων με αυτό τον προσανατολισμό.

2. Βλ., λ.χ., Putnam, H., «The meaning of "meaning"» στο Putnam, H., *Mind, Language and Reality*, Cambridge U. P. (1975), 215-71.

Θα μπορούσαμε ακόμη να ανακαλέσουμε το περίφημο αντιμηχανιστικό Gedankenexperiment του Leibniz:

17. Είναι κανείς εξάλλον υποχρεωμένος να δεχτεί ότι η αντίληψη και ο,τι εξαρτάται από αυτήν δεν εξηγείται με μηχανικούς λόγους, με σχήματα δηλαδή και κινήσεις. Κι αν υποθέσουμε ότι υπήρχε μηχανή, και ότι η δομή της προκαλούσε σκέψη, αίσθηση, αντίληψη, θα μπορούσε κανείς να τη φανταστεί σε μεγέθυνση, έτσι ώστε να διατηρούνται οι ίδιες αναλογίες, και να μπορεί κανείς να μπει μέσα σ' αυτήν, όπως σε μύλο. 'Εστω ότι δεχόμαστε κάτι τέτοιο' ο επισκέπτης δεν θα εύρισκε μέσα στη μηχανή παρά εξαρτήματα να ωθούν το ένα το άλλο, και τίποτε που να εξηγεί την αντίληψη. (Leibniz, G.W., «Les principes de la philosophie, dits Monadologie» (1714), στο τ.ί., *Oeuvres*, Παρίσι (1972), 398-9).

Δεν θα ήταν, νομίζω, αναχρονιστική υπερβολή να ανιχνεύσουμε εν σπέρματι στα κείμενα αυτά ιδέες όπως την περίφημη διαγνωστική «δοκιμασία» του Alan Turing για τη νοημοσύνη· τις ενστάσεις του John Searle ή του Ned Block κατά της τεχνητής νοημοσύνης ή του λειτουργισμού¹, και γενικότερα, πολλές από τις ενστάσεις κατά της αναγωγιστικής-μηχανιστικής προσέγγισης στη φιλοσοφία της διάνοιας.

Μια πρώτη παρατήρηση θα ήταν ότι πολλά από τα σύγχρονα επιχειρήματα, όσα στηρίζονται σε νοητικό πείραμα τα χαρακτηρίζει, ακριβώς όπως και τα κλασικά, έντονα και προγραμματικά απλουστευτική αντίληψη, τόσο για το τι είναι «μηχανισμός» ή «μηχανιστική εξήγηση» όσο και για τις ψυχοδιανοητικές λειτουργίες, όπως δείχνει η αντιπαραβολή με το γνωσιακό επιχείρημα. Ο Descartes και ο Leibniz, θα έλεγα, δεν θα δυσκολεύονταν να κατανοήσουν το άρθρο του Jackson άνευ ετέρου και να συμφωνήσουν με την επιχειρηματολογία του, ενώ θα χρειάζονταν πολλή προπαρασκευή για να παρακολουθήσουν οιοδήποτε κείμενο νευρο- ή ψυχο-φυσιολογίας. Οι πλασματικές καταστάσεις των νοητικών πειραμάτων είναι ίσως καλό προληπτικό μέσο κατά του «συνδρόμου των φιλοσόφων» (βλ. ανωτέρω)· ενδέχεται όμως ο δρόμος αυτός, σε συνδυασμό με την υπεραπλούστευση, να καταλήγει σε φιλοσοφικά άσκοπες και εμπειρικά άστοχες σχολαστικές υπερβολές, όποτε δεν σταθμίζονται δεόντως ουσιώδεις περιορισμοί που απορρέουν από τη φυσική και βιολογική συγκρότηση πραγματικών οργανισμών². Είναι τέλος κάπως οξύμωρο να

1. Βλ. Turing, A., «Computing Machinery and Intelligence», *Mind* LIX (1950), 433-60· Searle, J., «Minds, Brains, and Programs», *The Behavioral and Brain Sciences*, 3(1980), 417-24· Block, N., «Troubles with Functionalism», στο Block, N. (ed.), *Readings in the Philosophy of Psychology*, Harvard U. P., (198), τ. 1, 268-305.

2. Την πιο εμπεριστατωμένη, από όσο γνωρίζω, κριτική της κατάχρησης των ιστο-

θεωρούνται τα νοητικά πειράματα («κρίσιμα»), ικανά δηλαδή να διαιτητεύσουν τις διαφορές ανάμεσα σε αντίπαλες φιλοσοφικές αντιλήψεις, όταν μάλιστα ο αντιεπιστημονικός προσανατολισμός θα οδηγούσε τους δημιουργούς τους να αρνηθούν ότι είναι εφικτά κρίσιμα πειράματα ακόμη και στις εμπειρικές επιστήμες.

Το νοητικό πείραμα που θα εκθέσω στη συνέχεια έχει ρεαλιστικό χαρακτήρα, με την έννοια ότι δεν είναι πρακτικά ανέφικτο· από την άλλη, δεν το προτείνω ως λυδία λίθο φιλοσοφικών θέσεων - θα ήθελα απλώς να αναδείξει ορισμένες απορίες σχετικές με το χρωματικό βίωμα.

4. Η εμβριθής εγκεφαλοσκόπος μας δεν είναι απλώς κάτοχος κάθε μέλλουσας νευροφυσικής - έχει μελετήσει «πάντα» τα νευροεπιστημονικά. Δεν της έχει λοιπόν διαφύγει ότι είναι δυνατόν να βιώσει κανείς οπτικά αισθήματα ακόμη και όταν απουσιάζει εξωτερικό ηλεκτρομαγνητικό ερέθισμα με συχνότητα μέσα στα όρια ευαισθησίας του οπτικού οργάνου. Γνωρίζει δηλαδή ότι η αιτιακή ιστορία των συνειδητών οπτικών βιωμάτων δεν προϋποθέτει ότι ενεργοποιούνται όλα τα τμήματα της οπτικής οδού από τα φωτοευαίσθητα κύτταρα στον σμφιβληστροειδή έως τον οπτικό φλοιό στην ινιακή περιοχή του εγκεφάλου και προς τα ακόμη «υψηλότερα» κέντρα όπου επιτελείται η συνθετική ολοκλήρωση των εκρών από τα διάφορα υποσυστήματα. Πράγματι είναι δυνατό να υπάρξει οπτικό βίωμα με μόνη τη διέγερση του οπτικού φλοιού. Αυτά τα ενδογενή οπτικά φαινόμενα, τα φωσφάνεια, γνωστά ήδη από το 18ο αιώνα, έχουν συστηματικά μελετηθεί από το 1930 περίπου· περί το 1970 είχαν εκθρέψει ελπίδες ότι θα μπορούσε να κατασκευαστεί προσθετική συσκευή ικανή να αποδώσει, εν μέρει τουλάχιστον, την όραση σε όσους την είχαν στερηθεί από νόσημα, τραυματισμό ή χειρουργική επέμβαση¹.

ριών φιλοσοφικής φαντασίας στη φιλοσοφία της διάνοιας έχει διατυπώσει η Kathleen Wilkes στο έργο της *Real People*, Oxford U. P. (1988).

1. Βλ. για το ιστορικό και για μια πρώτη εισαγωγή, Oster, G., «Phosphenes», *Scientific American*, 222 (Φεβρ. 1970), 82-87· για τα φωσφάνεια πριν από κρίσεις ημικρανίας, βλ. Richards, W., «The Fortification Illusions of Migraines», *Scientific American*, 224 (Μάιος 1974), 88-97· δύο κλασικές νευροφυσιολογικές μελέτες σχετικές με την παραγωγή φωσφάνιων με άμεση ηλεκτρική διέγερση του οπτικού φλοιού είναι Brindley, G. S. και Lewin, W.S., «The Sensations Produced by Electrical Stimulation of the Visual Cortex», *J. Physiol.*, 196 (1968), 479-93· Dobbelle W. H. και Mladejovsky, M. G., «Phosphenes Produced by Electrical Stimulation of Human Occipital Cortex and their Application to the Development of a Prosthesis for the Blind», *J. Physiol.*, 243 (1974), 553-76. Πρβλ. Kellogg, R., Knoll, M. και Kugler, J., «Form-similarity between Phosphenes of Adults and Pre-school Children's Scribbling», *Nature*, 208 (Δεκ. 1965), 1129-30, όπου επισημαίνεται η μορφολογική ομοιότητα νηπιακών γραφήματων και φωσφάνειων.

Καθώς η αντίληψη της για τις γνωστικές διεργασίες ήταν ενεργητική, η Μαίρη δεν περίμενε να ελευθερωθεί και τα αισθητήρια της να καταγράψουν παθητικά τα κατάλληλα ερεθίσματα από τον έξω κόσμο. Αποκλείοντας βεβαίως το ενδεχόμενο να λάβει παραισθησιογόνα (LSD ή ψιλοκυβίνη) με έντονη χρωματογόνο δράση, η νευροφυσιολόγος μας επέλεξε να γίνει η ίδια το υποκειμένο πειραμάτων παραγωγής φωσφανείων με ηλεκτρική διέγερση του φλοιού του εγκέφαλου της. Η κατασκευή της συσκευής με τα απαιτούμενα κρανιακά ηλεκτρόδια και της γεννήτριες ηλεκτρικών παλμών δεν ήταν πέρα από τις δυνατότητες της. Τα σχετικά πειραματικά πρωτόκολλα ήταν σαφώς διατυπωμένα στην οικεία βιβλιογραφία και ευχερώς επαναλήψιμα. Πριν εκτελέσει το πείραμα, η Μαίρη είχε κάποια ανησυχία για το κατά πόσο θα της επέτρεπε όντως να υπερβεί τους περιορισμούς της αχρωματικής αγωγής της. Για ποιο λόγο; Ενώ σε όλες σχεδόν τις μελέτες αναφερόταν ότι εκτός από φωτεινά λευκά σχήματα παράγονται και έγχρωμα, και μάλιστα με όλες τις μεθόδους, ο πρώτος ασθενής του Brindley είχε τονίσει την απουσία χρωματικών αισθημάτων κατά την ηλεκτρική διέγερση του οπτικού φλοιού του.

Η ανησυχία όμως της Μαίρης αποδείχτηκε αβάσιμη. Η αισθητηριακή εμπειρία της κατά το πείραμα, συμφωνούσε με την περιγραφή των Dobelle και Mladejovsky:

«Οι ασθενείς Α και Β ανέφεραν μικρούς έγχρωμους ακτινωτούς τροχούς και ο ασθενής Γ τους περιέγραψε λεπτομερώς. Οι τροχοί μένουν σταθεροί στο οπτικό πεδίο, αλλά στρέφονται περί τον άξονα τους, με ακτίνες διαφόρων χρωμάτων».

Τα εν λόγω χρωματικά φαινόμενα είναι ζωηρά, και οι ασθενείς γενικώς αναφέρουν ότι τα φωσφάνεια είναι χρώματος ερυθρού, κινανού και ενίοτε πρασίνου. Θα είχε ενδιαφέρον να διατυπωθούν υποθέσεις για τη σχέση ανάμεσα στα χρώματα αυτά και στους βασικούς μηχανισμούς της χρωματικής αντίληψης» (Dobelle και Mladejovsky, ό.π., 368).

«Η επανάληψη της διέγερσης στο ίδιο σημείο συνήθως προκαλεί κάθε φορά το ίδιο αίσθημα. Την τρίτη μετεγχειρητική ημέρα, ο ασθενής Γ μπορούσε να εντοπίσει ποιά ηλεκτρόδιο διεγείρεται, χάρη στο σχήμα, στα ειδικά χρωματικά χαρακτηριστικά και τη θέση που παραγομένου φωσφανείου στο οπτικό πεδίο» (Dobelle και Mladejovsky, ό.π., 370).

Πώς θα περιγράφαμε τη γνωσιολογική κατάσταση της Μαίρης, πριν από την απελευθέρωση της, αλλά μετά την ευτυχή έκβαση του πειράματος; Πρώτα πρώτα, η Μαίρη, χάρη στην πλήρη θεωρητική και πρακτική νευροφυσιολογική γνώση της, κατόρθωσε να αποκτήσει ό,τι, κατά το γνωσιακό επιχείρημα, της έλειπε: χρωματικό βίωμα. Επομένως, κατά το γνωσιακό επιχείρημα πάντα, η Μαίρη, χωρίς να εγκαταλείψει το αχρωμα-

τικό δωμάτιο της, γνωρίζει πλέον ο,τι μόνη η νευροφυσιολογία δεν της είχε μάθει (ακριβώς όπως αν ήταν ηρωίδα μιας άλλης ιστορίας, η Μαίρη, εμβριθής μαιευτήρας, θα μάθαινε πώς είναι οι ωδίνες μόνο όταν θα γινόταν μητέρα η ίδια). Το ζήτημα, νομίζω, διακλαδίζεται προς δύο κατευθύνσεις: αφενός, πώς εξηγείται το οπτικό φαινόμενο των φωσφαινών, και αφετέρου, ποια είναι η νέα γνωσιολογική κατάσταση της Μαίρης. Ας δούμε κάθε ομάδα προβλημάτων με τη σειρά.

Το νοητικό πείραμα μας, αφορά ένα φαινόμενο όπου, κατά κάποιον τρόπο «βραχυκυκλώνεται» το τμήμα της οπτικής οδού από τους περιφερειακούς φωτοεαίσθητους υποδοχείς έως την οικεία περιοχή του φλοιού. Επομένως το φωτοαίσθημα στην περίπτωση αυτή παράγεται ανεξάρτητα από οιαδήποτε διέγερση των περιφερειακών οργάνων, και άρα δεν χρειάζεται να αντιμετωπιστεί το κλασικό, τουλάχιστον από την εποχή το Locke, αν όχι του Δημόκριτου, ζήτημα της οντολογίας των χρωμάτων: τα χρώματα είναι «αντικειμενικά» ή «υποκειμενικά»; είναι «πρωτεύουσες» ή «δευτερεύουσες» ποιότητες; πρόκειται για φυσικές ιδιότητες της επιφάνειας των αντικειμένων ή παράγονται «μέσα στο κεφάλι»; Τα φωσφαινεία, μαζί με τις ψευδαισθήσεις, τις παραισθήσεις, τα μεταισθήματα, τα μεταμερή, μοιάζουν να ωθούν τα χρωματικά αισθήματα προς τον πόλο του υποκειμενικού¹. Αν τώρα, τα αισθήματα αυτά, ως «δεδομένα-των-αισθήσεων», πρέπει να θεωρηθούν υποστάσεις, επέκεινα του φυσικού κόσμου, είναι μια άλλη ιστορία. Αυτή ακριβώς είναι η επιλογή του δημιουργού της Μαίρης, του Frank Jackson, στη μελέτη του *Perception* (Cambridge U.P. (1977)) - το γνωσιακό επιχείρημα είναι απλώς συμπυκνωμένη έκφραση της βασικής θέσης της μελέτης αυτής. Ο David Hilbert² αποφεύγει τα μεταφυσικά προβλήματα, όσα αναπτύττον σχετικά με τα «δεδομένα-των-αισθήσεων», και ταυτίζει τα χρώματα με ένα αντικειμενικό γνώρισμα, τη φασματική ανακλαστικότητα επιφανειών (spectral surface reflectance). Ωστόσο η θεωρία αυτή αδυνατεί να εξηγήσει αφενός τα φαινόμενα της κατηγορίας των ψευδαισθήσεων και αφετέρου ουσιώδη γνωρί-

1. Σχετικά με την ψυχοφυσιολογία της έγχρωμης όρασης, πολύ προσιτή και ενδιαφέρουσα εισαγωγή βρίσκουμε στο κεφ. 6: «Color Perception» στο Sekuler, R. και Blake, R. (eds), *Perception*, N.Y. (1994³), 172-207· συντομότατη εισαγωγή στη νευροφυσιολογία της όρασης στο Despopoulos, A. και Silbernagl, S., *Εγχειρίδιο Φυσιολογίας*, Αθήνα (1989), 300-15· και εμπειρικά πολύ ενήμερη και φιλοσοφικά κατατοπιστική πραγματεύση, βλ. Hardin, C.L.[1], *Color for Philosophers. Unweaving the Rainbow*, Indianapolis, (1988/1993)· για αντίλογο στην προσέγγιση του Hardin, βλ. van Brakel, J., «The Plasticity of Categories», *British Journal for the Philosophy of Science*, 44 (1993), 103-35 και την απάντηση του Hardin, «Van Brakel and the Not-so-Naked Emperor», *British Journal for the Philosophy of Science*, 44 (1993), 137-50.

2. Hilbert, D., *Color and Color Perception*, Univ. of Chicago Press, (1987).

σματα των χρωμάτων, όπως ότι τα «βασικά» χρώματα ανά δύο βρίσκονται σε σχέση «αντίθεσης»: όταν η νευρονική διεργασία που υποτείνει το αίσθημα του ερυθρού γίνεται εντονότερη, εξασθενεί συνακόλουθα εκείνη του πράσινου, και ανάλογα για το κυανό και το κίτρινο¹.

Η θεωρία του C.L. Hardin διακρίνεται κατά το ότι είναι την ίδια στιγμή «υποκειμενική» σε ό,τι αφορά τα χρώματα και αναγνωριστικά υλιστική σε ό,τι αφορά το χρωματικό βίωμα². Με αφορμή τη θεωρία του Jackson για την αισθητηριακή αντίληψη, και την παράλειψη των υλιστών να αποδώσουν τη δέουσα προσοχή στα ενδογενή χρωματικά αισθήματα, ο Hardin (Hardin [1], 91-112) δείχνει ότι όσοι δέχονται τον υποκειμενικό χαρακτήρα των χρωμάτων δεν δεσμεύονται κατ' ανάγκη να δεχτούν τα χρώματα ως «δεδομένα-των-αισθήσεων» - φαινόμενα μέσα σε ένα μη υλικό «οπτικό πεδίο»: η υποκειμενική θεωρία συμβιβάζεται κάλλιστα με τα χρώματα ως «νευρονικά συμβάντα». Την τελευταία αυτή θέση ο Hardin τη στηρίζει στην ανάλυση και στην εξήγηση ψευδαισθήσεων. Μόνο μία αφορά χρωματικό βίωμα (ψευδαισθηση von Bezold) - δύο έχουν σχέση με την αναγνώριση γραμμικών σχημάτων (οφθαλμαπάτη Müller-Lyer, κύβος Necker) και η άλλη, με το φαινόμενο Mach, όπου η φωτεινότητα επιφάνειας φαίνεται να αλλάζει ανάλογα με το πώς ο θεατής αντιλαμβάνεται το περίγραμμα. Επιμένω, με κίνδυνο να βάλω σε δοκιμασία την υπομονή όσων με έχουν ακολουθήσει έως εδώ, γιατί θεωρώ σημαντικό ότι ο Hardin στηρίζει τη θέση του σε εξονυχιστική ανάλυση επιμέρους φαινομένων και όχι σε ευρείες γενικεύσεις που απαλείφουν ex ante σημαντικά γνωρίσματα της εμπειρικής πραγματικότητας. Σε ό,τι αφορά την όραση (αλλά και όλες εν γένει τις διανοητικές λειτουργίες, από τις απλούστερες έως τις πιο ολοκληρωμένες) πρέπει να τονιστεί ότι από πολλές πλευρές τεκμηριώνεται³ η εξής εικόνα: σχετικά αυτόνομοι, ιεραρχι-

1. Η αντιθετική σχέση των χρωμάτων ανά ζεύγη είναι η βασική παραδοχή στο ευρύτατα διαδεδομένο σήμερα θεωρητικό υπόδειγμα για τη χρωματική όραση: βλ., λ. χ., Hurvich, L., Jameson, D., «An opponent-process theory of color vision», *Psychological Review*, 64 (1957), 384-408. Δεν έχω εντυπώσει στο κατά πόσον αντικειμενικές θεωρίες για τα χρώματα, όπως του Hilbert, εναρμονίζονται με το άλλο δημοφιλές υπόδειγμα, τη «θεωρία Retinex» του E. Land - βλ. Land, E. H., «The retinex theory of color vision», *Scientific American*, 237 (Ιούν. 1977), 108-28-, έχω όμως λόγους να πιστεύω ότι ίσως υπάρχουν προβλήματα, δεδομένου του υποκειμενικού χαρακτήρα της θεωρίας retinex.

2. Στο σημαντικό άρθρο τους «Physicalist Theories of Color», *Philosophical Review*, C (1991), 67-106, οι P. A. Boghossian και J. D. Velleman, ονομάζουν «φυσικοκρατικές» όσες θεωρίες αποδίδουν τα χρώματα στα αντικείμενα - και προτείνουν εκτενή φιλοσοφική ανασκευή τους.

3. Σχετικά με τις επιμέρους λειτουργίες που η ολοκλήρωσή τους παράγει την όραση, ενδείξεις μπορεί να συλλέξει κανείς στο Sekuler και Blake, ό.π., όπου εκτίθεται, π.

νά διασυνδεδεμένοι ανατομολειτουργικοί «σπόνδυλοι» (modules) στο κεντρικό νευρικό σύστημα επιτελούν ο καθένας πολύ εξειδικευμένο έργο- άλλοι επεξεργάζονται τα γραμμικά σήματα (σχήματα, περιγράμματα), άλλοι τα κινούμενα σχήματα, άλλοι τα χρώματα, άλλοι συνθέτουν στοιχεία (λ.χ., χρώμα-κίνηση), ώστε δεν είναι πάντοτε φρόνιμο να συμφύρονται στις ίδιες αδρομερείς φιλοσοφικές κλάσεις (λ.χ., αίσθηση) φαινόμενα με πολύ διαφορετική ανατομολειτουργική υφή και ιστορία.

Έχοντας κάνει τις επισημάνσεις αυτές, πρέπει, για χάρη της συνέπειας, να ομολογήσω ότι, επί του παρόντος και όσο γνωρίζω δεν υπάρχει ικανοποιητική νευροφυσιολογική εξήγηση των φωσφαιίων, ανάλογη προς εκείνη των φαινομένων που μόλις προανέφερα. Ούτε μπορώ να είμαι βέβαιος πώς η θεωρία της αντίθεσης των χρωμάτων, ή κάποια άλλη ψυχοφυσιολογική θεωρία, θα κάλυπτε επαρκώς το φαινόμενο, παρόλη τη σχετικά απλή δομή του-απλή, με την έννοια ότι οι όροι παραγωγής του περιλαμβάνουν λίγα βήματα, καθώς δεν εμπλέκεται ολόκληρη η οπτική οδός.

Ωστόσο, θα αποπειραθώ να σκιαγραφήσω προς ποια κατεύθυνση θα μπορούσε κανείς να αναζητήσει εξήγηση. Οιαδήποτε σχετικά πλήρης θεωρητική πραγμάτευση θα έπρεπε να καλύπτει τουλάχιστον τα βασικά φαινομενολογικά χαρακτηριστικά των φωσφαιίων: το σχήμα (ακτινωτός τροχός), τη χροιά, την κίνηση (περιστροφή γύρω από τον άξονα), και το ότι βιώνονται ως προβαλλόμενα μπροστά από τα μάτια σε απόσταση περίπου 30 έως 40 cm. Πρώτα πρώτα, να παρατηρήσουμε ότι η Μαίρη δεν χρειάζεται μια «ιδιωτική» γλώσσα προκειμένου μπορεί να αναφέρεται στο νέο της βίωμα. Αν θεωρηθεί ότι τα φωσφαιία είναι σταθερά και αναπαραγώγιμα κατά βούληση (βλ. σ. 113), η Μαίρη θα μπορούσε να κατασκευάσει ονοματοειδείς «άκαμπτους σημειωτήρες» - κατ' αναλογία προς ο,τι ο Kripke ονομάζει «rigid designators»¹. Για παράδειγμα, θα μπορούσε να αναφέρεται στον «ακτινωτό κύκλο στη θέση B4 του οπτικού πεδίου μου, όταν είναι αχώγιμο το ηλεκτρόδιο αριθ. 40, το πλάτος του ηλεκτρικού παλμού 3 mA, η διάρκεια του παλμού 0,25 msec».

χ., ερευνητικά πορίσματα των Hubel και Wiesel, Zeki, Ramachandran. Για το σπονδυλωτό χαρακτήρα των υπόλοιπων διανοητικών λειτουργιών, μια πρώτη αντίληψη μπορεί κανείς να σχηματίσει διατρέχοντας το Glass, A. L. και Holyoak, K. J., *Cognition*, (1986) - οι σχετικές μελέτες βασίζονται στην επιλεκτική απώλεια ικανοτήτων που συνοδεύει εντοπισμένες νευρολογικές βλάβες. Τα σχετικά δεδομένα έχουν μεγάλη σημασία για το γενικότερο προβληματισμό γύρω από τη φύση και τη συγκρότηση του συνειδητού βιώματος.

1. Βλ., Kripke, S., *Naming and Necessity*, Οξφόρδη (1980) - ίσως οι «άκαμπτες περιγραφές» της Μαίρης να είναι απλώς *de facto* και όχι *de jure* (21). Δεν είμαι βέβαιος κατά πόσο ο Kripke, με τον αντιυλιστικό προσανατολισμό του, θα ήταν ευτυχής με μια τέτοια χρήση των ιδεών του.

Δεδομένου εξάλλου ότι τα ηλεκτρόδια είναι τοποθετημένα σε καλώς ορισμένα σημεία του κρανιακού χάρτη και είναι γνωστό ποια ηλεκτρόδια συσχετίζονται με τη διέγερση συγκεκριμένων σημείων προβολής του (φαινομενολογικού-ψυχολογικού) οπτικού πεδίου στον οπτικό φλοιό, οι «άκαμπτοι σημειωτήρες» θα μπορούσαν να διεισδύουν έως το επίπεδο των νευρονικών συμβάντων, με τοπογραφική διακριτική ικανότητα επί του φλοιού μεταξύ 2 και 3 mm (Dobelle και Mladejovsky, *ό.π.*, 560, 571). Όσο για τα επιμέρους φαινομενολογικά χαρακτηριστικά των φωσφαινίων θα ήταν εφικτό να συσχετιστούν με λειτουργικά εξειδικευμένες ομάδες νευρώνων του οπτικού φλοιού. Έχουν πράγματι εντοπιστεί ποιές νευρονικές ομάδες εξειδικεύονται στην επεξεργασία των σχημάτων, των χρωμάτων και της κίνησης¹. Το κατά πόσο μπορεί να διαπιστωθεί αν, στο συγκεκριμένο πειραματικό πλαίσιο, διεγείρονται οι οικείες νευρονικές ομάδες, όσες συσχετίζονται με τα αντίστοιχα φαινομενολογικά χαρακτηριστικά, εξαρτάται από την ανατομολειτουργική διακριτική ικανότητα των μεθόδων μας. Οι όποιες δυσκολίες είναι πρακτικές, όχι αρχής. Θα άξιζε τέλος να τονιστεί ότι οι «άκαμπτοι σημειωτήρες» μας σώζονται από το νυστέρι του επιχειρήματος του Wittgenstein για την ιδιωτική γλώσσα, καθώς η ταύτιση του κατονομαζομένου γίνεται μέσω «δημοσίων» παραμέτρων. Επιπλέον, αυτός ο τρόπος αναφοράς στο βίωμα είναι οντολογικά ουδέτερος - «topic-neutral», κατά τους Smart και Armstrong - δεν συνεπάγεται ότι συμβαίνει κάτι φυσικό (αφού μνημονεύει το φαινόμενο βίωμα) ούτε όμως και ότι συμβαίνει κάτι μη φυσικό². Θα μπορούσε κανείς βεβαίως να διερωτηθεί αν με την απόπειρα αυτή για «φυσικρατική» εξήγηση των φωσφαινίων, ή μάλλον για σταδιακή και λεπτομερή συσχέτιση των φαινομενολογικών χαρακτηριστικών τους με την υποκείμενη νευροφυσιολογία, γεφυρώνεται το περιβόητο «εξηγητικό χάσμα». Η φρονιμότερη ίσως απάντηση θα ήταν αυτή του Hardin: «...at the present rudimentary state of our knowledge of the visual system, most of this is promise, program, and principle».

Ας δούμε όμως τώρα από πιο κοντά την εξέλιξη στη γνωσιολογική κατάσταση της Μαίρης. Τι ακριβώς απ' όλα όσα γνωρίζει είναι καθορι-

1. Βλ., π.χ., Hubel D. H., και Wiesel, T. N., «Brain Mechanisms of Vision», *Scientific American*, 241 (1979), 150-63· Ramachandran, V. S., και Anstis, S. M., «The perception of apparent motion», *Scientific American*, 254 (1986), 102-9· Zeki, S. M., «The representation of colours in the cerebral cortex», *Nature*, 284 (1980), 412-8. Για ολοκληρωμένη εικόνα των ερευνητικών πορισμάτων και του ιστορικού πλαισίου, βλ. το *opus magnum* του Semir Zeki, *A Vision of the Brain*, Οξφόρδη (1993).

2. Βλ. Armstrong, D. M., *A Materialist Theory of the Mind*, Λονδίνο 1896, 121.

στικό εδώ; Οπωσδήποτε κατέχει ένα είδος χρωματικού λεξικού όπου συσχετίζονται από τη μια, τα φυσικά χαρακτηριστικά - λ.χ., συχνότητα και ένταση - των χρωματικών ερεθισμάτων και από την άλλη τα ψυχοφυσιολογικά χαρακτηριστικά του βιώματος - λ.χ., χροιά (hue), λαμπρότητα (brightness), κορεσμός (saturation). Γνωρίζει επίσης ότι έχει δει χρώματα. Δεν γνωρίζει όμως πιο έγχρωμο σχήμα στο οπτικό πεδίο της είναι «κόκκινο», ποιο «πράσινο», ποιο «κυανό»: κατονομάζει τα έγχρωμα σχήματα μέσω των «άκαμπτων σημειωτήρων» της και αγνοεί ποιός από αυτούς είναι συνώνυμος - έχει την ίδια αναφορά - με τον όρο «κόκκινο», κλπ. Η «γνώση» της είναι κατ' ουσίαν συντακτική, η σημασιολογική συνιστώσα του λεξιλογίου της μόνο φαινομενική. Γνωρίζει την ένταση όσων όρων χρησιμοποιεί για να δηλώσει τις εντυπώσεις της, όχι όμως την έκτασή τους: σε ποια αντικείμενα, εκτός από τις εντυπώσεις της, θα μπορούσαν να αποδοθούν αληθώς οι όροι αυτοί. Δεν νομίζω να μπορούσε καν να γίνει εδώ λόγος για εντασιακό ισομορφισμό (intensional isomorphism)¹ ανάμεσα στο λεξιλόγιο των χρωμάτων και στους «άκαμπτους σημειωτήρες» των εντυπώσεων. Η ψευδοσημασιολογική φύση του λεξιλογίου της εξηγείται από το ότι τα χρωματικά αισθήματα της δεν προέκυψαν μέσα από τη φυσιολογική αιτιακή οδό - με τη διέγερση δηλαδή των περιφερειακών φωτούποδοχέων, στη διεπαφή του εξωτερικού και του διανοητικού κόσμου.

Όσο η εμβριθής εγκεφαλοσκόπος θα «γνώριζε» μόνο τις εντυπώσεις της, χωρίς να μπορεί να τις ενσωματώσει σε ένα αρκετά πλούσιο πλέγμα από πεποιθήσεις, προβλέψεις, προσδοκίες, ικανό να στηρίξει τη σκόπιμη δράση μέσα στον κόσμο, είναι αμφίβολο αν αυτή η ομφαλοσκοπική παράσταση θα άξιζε όντως να ονομαστεί «γνώση». Κι αυτό όχι γιατί το αντικείμενο της παράστασης, τα ποιοτικά γνωρίσματα του βιώματος, είναι «επιφαινόμενα» και είτε υπάρχουν είτε δεν υπάρχουν δεν αλλάζουν τίποτα στο φυσικό κόσμο, αλλά γιατί η συρρίκνωση του συνειδητού βιώματος σε ένα παθητικό «χώρο ποιότητων», στερεί προγραμματικά το γνωστικό υποκείμενο από τα μέσα αλληλεπίδρασης με τον κόσμο.

Από την άλλη, η ηρώιδα μας θα βελτίωνε τη γνωστική κατάσταση της, αν κατόρθωνε να συγκρίνει τα βιώματα της με τα αντίστοιχα των άλλων. Θα μπορούσε να αποκαταστήσει έμμεσα το πλέγμα γνήσιας σημασιολογικής αναφοράς του λεξιλογίου της, αν έπειθε τη Μάρθα - τη φίλη της στον «έξω κόσμο» - να υποστεί την ίδια ακριβώς πειραματική δοκιμασία. Έστω ότι η έκβαση της δοκιμασίας ήταν σύμφωνη με τις επιστημονικές προβλέψεις και τις ανθρώπινες προσδοκίες της Μάρθης: η Μάρ-

1. Βλ. Carnap, R., *Meaning and Necessity*, Σικάγο (1956), 59.

θα θα είχε πράγματι δει ένα «ακτινωτό κύκλο στη θέση Β4...» όπως και η Μαίρη. Η Μάρθα όμως, έχοντας ζήσει πάντα στον έξω κόσμο, θα μπορούσε να προσθέσει «ο κύκλος αυτός έχει χρώμα κόκκινο, όπως οι ώριμες τομάτες». Όντας καλή φίλη, θα μετάφερε την πληροφορία αυτή στη Μαίρη. Και η Μαίρη, βέβαιη για την ειλικρίνεια της Μάρθας¹, θα έκανε τότε το πρώτο βήμα της στον πραγματικό κόσμο - στο φυσικό μαζί και ανθρώπινο κόσμο των χρωμάτων!

33. Για να αντιμετωπίσει τους σκεπτικιστές, η Μαίρη θα έπρεπε να τοποθετηθεί και στη γνωστή ένσταση τη στηριγμένη στο «αντεστραμμένο φάσμα» (inverted spectrum). Θα έπρεπε δηλαδή να αποκλείσει ότι η Μάρθα εμφανίζει τη συγγενή ιδιομορφία να βλέπει λ.χ., κόκκινο ο,τι οι άλλοι βλέπουν πράσινο. Αν η Μάρθα είχε το χαρακτηριστικό αυτό, ονόμαζε συστηματικά «κόκκινο» ο,τι οι άλλοι βλέπουν πράσινο και αντιστρόφως. Κατά τους αντιυλιστές, η αντιστροφή αυτή βιώνεται «σε πρώτο πρόσωπο» και δεν διαπιστώσιμη από τη σκοπιά «του τρίτου προσώπου». Η Μαίρη θα μπορούσε να δεχτεί ότι τη φίλη της απλά δεν την χαρακτηρίζει η αντιστροφή αυτή, αιτιολογώντας ίσως την άποψη της στατιστικά: επικαλούμενη δηλαδή τη σπανιότητα της περίπτωσης. Βεβαίως δεν θα ανασκεύαζε έτσι την ένσταση. Ωστόσο, ως έμβριθης νευροεπιστήμων θα γνώριζε ότι το αντεστραμμένο φάσμα είναι εννοιολογικά ενδιαφέρον ενδεχόμενο, που όμως μπορεί να αποκλειστεί με λίγη περισσότερη προσοχή στις πραγματικές λεπτομέρειες της χρωματικής δράσης. Είναι όντως γνωστό ότι τα χρώματα, εκτός από τη χροιά, τον κορεσμό και τη λαμπρότητα, χαρακτηρίζονται και ως «θερμά» - π.χ., το κόκκινο - και «ψυχρά» - π.χ., το πράσινο. Τα θερμά χρώματα προκαλούν μετρήσιμη αύξηση των καρδιακών παλμών ή μετρήσιμη μεταβολή στην επιδερμική αγωγιμότητα χάρη στην έκκριση ιδρώτος - στα δεδομένα αυτά στηρίζεται, λ.χ., η «ψυχολογική» χρήση χρωμάτων σε νοσοκομεία, σχολεία, χώρους εργασίας ή αναψυχής. Η Μαίρη λοιπόν θα είχε τρόπο να διαπιστώσει αν ο,τι κατονομάζει η Μάρθα αντιστοιχεί σε θερμό ή ψυχρό χρώμα. Για το αντεστραμμένο φάσμα, μόνο ενδεικτικά, βλ., Lycan, W.G., «Inverted Spectrum», *Ratio*, 15 (1973), 315-19· Shoemaker, S., «The Inverted Spectrum», *Journal of Philosophy*, LXXIX (1982), 357-81. Horgan, T., «Functionalism, Qualia, and the Inverted Spectrum», *Philosophy and Phenomenological Research*, XLIV (1984), 453-69· η ανησυχία μήπως πρόκειται για ιδέες του συρμού, διασκεδάζεται εύκολα με την ανάγνωση του *Essay Concerning Human Understanding*, του Locke (1690).

G. MARAGOS

WHAT COULD THE COMPLEAT AUTOCEREROSCOPIST LEARN?

In his widely discussed paper «Epiphenomenal Qualia», F. Jackson imagines Mary living since her birth in an exclusively black-and-white environment and communicating with the outside world by means of her black-and-white television set. Eventually Mary manages to master all neurophysiological knowledge concerning *inter allia* colour vision. Upon her release, Mary learns what it is like to see a red tomato or the blue skies, experiencing the relevant sensations. Therefore, her knowledge prior to her release was not complete. According to Jackson this shows that materialism is false, since at least some mental phenomena, in this case the *qualia* of sensation, are not captured by neurophysiology. *Pace* Jackson the *qualia* of sensation, being *epiphenomena*, are causally inert and hence not detectable by objective, scientific, means.

In this paper we set out to show that knowing absolutely everything about colour vision might be more than Jackson allows. Prior to her release, while still in her black-and-white room, Mary could arrange to be the subject in an experiment leading to her experiencing colour sensations, *phosphenes*, by direct electrical stimulation of her visual cortex. Such endogenous colour experiences are to a certain extent repeatable and controllable so that she can refer to them by quasi rigid designators, avoiding thus the pitfalls of the private language argument. Mary's epistemic situation after the experiment is such that she has the long sought for sensations, but it is doubtful whether she now possesses *knowledge* of the sort that would invalidate the claims of materialism. Mary could leverage herself to something like knowledge, that is a system of beliefs apt to sustain purposeful activity in the real world, ...with a little help from her friends.