

Η αναγωγή του ψυχολογικού στο νευροβιολογικό και γιατί τη φοβόμαστε

(Το πρόβλημα της αναγωγής στην αναπαραστατική θεωρία του νου και της συνείδησης)

Βαγγέλης Καφετζόπουλος, Ξυλόκαστρο 2005

Το ερώτημα αφετηρίας¹

Έχει η έρευνα στις Νευροεπιστήμες φέρει στο φως τεκμήρια που θα μπορούσαν να πείσουν τον όποιον *λογικό και καλοπροαίρετο άνθρωπο, τον Σωκράτη, ας πούμε, πως όπως ακριβώς η λύρα παράγει μουσική έτσι κι ο εγκέφαλός του παρήγαγε τη χαρά του για το επικείμενο ταξίδι του στα λημέρια των Θεών, η οποία χαρά, παρεμπιπτόντως, θα εξανεμίζονταν τη στιγμή που θα αποδέχονταν την παραπάνω πρόταση;*

Γιατί χαιρέται ο Σωκράτης;

Μια πρώτη απάντηση από τη σκοπιά του *λογικού και καλοπροαίρετου* «αναγωγιστή» νευροεπιστήμονα θα ήταν ότι η χαρά του Σωκράτη οφείλεται στην ενεργοποίηση του συστήματος ανταμοιβής (reward system) του εγκεφάλου του, που με τη σειρά της ενεργοποιεί τις συγκινησιακές αντιδράσεις του οργανισμού. Από την ανακάλυψή του στον εγκέφαλο του επίμυ (Olds and Milner 1954), το σύστημα ανταμοιβής αποδείχτηκε ότι συμμετέχει στην γένεση ευχάριστων συναισθημάτων στον άνθρωπο (Heath 1964), ενώ μέσω του ενισχυτικού του ρόλου φαίνεται να εμπλέκεται και σε πιο σύνθετες συμπεριφορές, όπως στη μνήμη και τη μάθηση () ή την εξάρτηση από εξαρτησιογόνες ουσίες (Koob and Le Moal 1997).

¹ Η χρησιμοποίηση της πρότασης «*όπως η λύρα παράγει μουσική*» αντί της ιστορικής «*όπως το συκώτι παράγει τη χολή*» θέτει ορισμένα μεθοδολογικά και φιλοσοφικά προβλήματα, επειδή καμία φυσική (εμπειρική) επιστήμη δεν φαίνεται να μπορεί να απαντήσει ακόμη και στο πρώτο σκέλος, πώς η λύρα παράγει μουσική. Η λύρα παράγει κύματα πίεσης του αέρα, που γίνονται ήχοι μόνον όταν κάποιος οργανισμός, με ανεπτυγμένο ακουστικό σύστημα, βρίσκεται κοντά για να την ακούσει. Τα κύματα του αέρα μετατρέπονται σε ήχο μόνο μέσα στον εγκέφαλο του υποκειμένου που τα αντιλαμβάνεται και αυτή τη μετατροπή τη μελετά η φυσιολογία. Σε μουσική μετατρέπονται στους εγκεφάλους συγκεκριμένων οργανισμών (π.χ. στον άνθρωπο και πιθανόν όχι στα ζώα), που έχουν μάθει να ακούν τους συγκεκριμένους ήχους ως μουσική. Εάν ένας άνθρωπος δεν έχει ακούσει ποτέ μουσική μάλλον δεν θα αντιληφθεί αυτήν την ακολουθία των ήχων ως μουσική, όπως ένας άνθρωπος που δεν έχει μεγαλώσει στο κατάλληλο γλωσσικό περιβάλλον ποτέ δεν θα αντιληφθεί τους γλωσσικούς ήχους ως γλώσσα (όταν π.χ. ακούμε κινέζικα, ακούμε ήχους και όχι γλώσσα). Μια διευρυμένη νευροεπιστήμη που θα περιελάμβανε και την «ψυχο-μουσικολογία» (κατ' αντιστοιχία με την ψυχο-γλωσσολογία) ίσως να μπορούσε να μας πει πώς γίνεται αυτή η κατανόηση της μουσικής. Καμία εμπειρική επιστήμη όμως δεν μπορεί να μας πει γιατί ο ένας λυράρης «*παίζει καλύτερα*» από τον άλλον, γιατί αυτή η μουσική είναι «*πιο ωραία*» από εκείνη. Κι επειδή η *ουσία της μουσικής είναι η ομορφιά της* (η αισθητική της αξία), *καμία επιστήμη* δεν μπορεί να την εξηγήσει γιατί *καμία επιστήμη δεν μπορεί να επικυρώσει λογικά τις αισθητικές (όπως άλλωστε και τις ηθικές) μας κρίσεις*. Για την οικονομία αυτής της παρουσίασης, όμως, θα υποθέσουμε ότι η παραγωγή της μουσικής είναι ένα φυσικό φαινόμενο και θα προσπαθήσουμε να δείξουμε ότι και η χαρά που νιώθει ο Σωκράτης είναι ένα φυσικό φαινόμενο, το οποίο όμως δεν ανάγεται πλήρως στη λειτουργία του *δικού του* εγκεφάλου, όπως η αισθητική αξία της μουσικής δεν ανάγεται πλήρως στη *συγκεκριμένη λύρα* που την παίζει.

Μια τέτοιου τύπου άμεση αναγωγή μιας ψυχολογικής κατάστασης σε έναν εγκεφαλικό μηχανισμό, αν και μπορεί να εγείρει ενστάσεις ως προς ορισμένα μεθοδολογικά προβλήματα, αφήνει να διαφανεί ένα ακόμη σημαντικό πρόβλημα στη θέση του ερωτήματος αφετηρίας. Πρώτον, η αναγωγή μιας συγκινησιακής κατάστασης σε εγκεφαλικές καταστάσεις είναι μάλλον εύκολη σε σχέση με την αναγωγή νοητικών ή συνειδησιακών καταστάσεων. Δεύτερον, λόγω της φύσης των συγκινησιακών καταστάσεων, που περιλαμβάνουν αντιδράσεις του αυτόνομου νευρικού συστήματος (ΑΝΣ), μια τέτοια αναγωγή θα μπορούσε να είναι «εξαλειπτικού» χαρακτήρα, να ερμήνευε δηλαδή εξ ολοκλήρου τη χαρά στο εγκεφαλικό επίπεδο έτσι ώστε οι ψυχολογικές ερμηνείες να εξαλείφονταν. Τρίτον, αυτή η εξάλειψη αφορά τις «ψυχολογικές ερμηνείες» της χαράς και όχι την ίδια τη χαρά. Η χαρά δεν εξαλείφεται ως υποκειμενική εμπειρία (εφόσον η αναγωγή αφορά τη χαρά ως συγκινησιακή κατάσταση) άρα δεν υπάρχει περίπτωση να εξανεμισθεί, όπως τίθεται στο ερώτημα αφετηρίας. Ο «φόβος της αναγωγής» πιθανόν να οφείλεται σε αυτήν την παρανόηση, την οποία θα προσπαθήσουμε να διαλευκάνουμε και στη συνέχεια.

Οι νοητικές αναπαραστάσεις και οι ιδιότητές τους

Στην περίπτωση της χαράς ως συγκινησιακής κατάστασης ένας από τους λόγους της ευκολίας της αναγωγής στις εγκεφαλικές καταστάσεις είναι η ύπαρξη μιας κοινής γλώσσας περιγραφής των δύο επιπέδων, εφόσον, όπως είπαμε, οι συγκινησιακές καταστάσεις μπορούν να περιγραφούν με όρους αντιδράσεων του ΑΝΣ.² Στην περίπτωση όμως νοητικών ή συνειδησιακών καταστάσεων η κοινή γλώσσα δεν είναι δεδομένη και πρέπει να αναζητηθεί. Μια καλή αφετηρία λοιπόν για τη δυνατότητα της αναγωγής αυτών των καταστάσεων σε εγκεφαλικές καταστάσεις θα ήταν η διερεύνησή της μέσω της *αναπαραστατικής θεωρίας του νου*³, η οποία αποτελεί το κοινό δόγμα της γνωσιακής επιστήμης. Η *γνωσιακή επιστήμη* συγκροτείται ως *κοινό ερευνητικό πρόγραμμα* τεσσάρων βασικών πεδίων, της γνωστικής ψυχολογίας, των νευροεπιστημών, της επιστήμης των ηλεκτρονικών υπολογιστών και της φιλοσοφίας του νου και ένα από τα βασικά αιτήματά της είναι ακριβώς η αναζήτηση κοινής γλώσσας περιγραφής των νοητικών φαινομένων στα επιμέρους επιστημονικά πεδία. Ένα επιπλέον πλεονέκτημα της αναπαραστατικής θεωρίας του νου είναι ότι πρόκειται για μια καλά δομημένη *φυσικαλιστική* (υλιστική) θεωρία των νοητικών φαινομένων, γεγονός που διευκολύνει τη διερεύνηση της δυνατότητας αναγωγής τους ή μη σε εγκεφαλικά φαινόμενα.

Πώς τίθενται λοιπόν οι όροι του ερωτήματος αφετηρίας με τους όρους της αναπαραστατικής θεωρίας του νου;

- Ο Σωκράτης **χαίρεται** επειδή **πιστεύει ότι** θα πάει στα λημέρια του θεού του.
- Ο θεός του Σωκράτη –ο Δίας– δεν υπάρχει (κι αυτό ίσως το γνώριζε και ο ίδιος).

² Και στις δύο επικρατούσες σήμερα ψυχολογικές θεωρίες του συναισθήματος οι σωματικές συνιστώσες είναι σημαντικές και προηγούνται της συναισθηματικής εμπειρίας. Στη μεν θεωρία James-Lange (βλ. Lange and James 1922) η συναισθηματική εμπειρία είναι η *αντίληψη* των σωματικών αλλαγών που προκαλούν περιφερικά ή εσωτερικά ερεθίσματα, στη δε θεωρία του Schachter (Schachter and Singer 1962) είναι η *γνωστική ερμηνεία* τους.

³ Η *αναπαραστατική θεωρία του νου* αναφέρεται συχνά και ως *αναπαραστατική θεωρία της συνείδησης* (π.χ. βλ. Lycan 2005). Οι δύο όροι φαίνεται να είναι εναλλακτικοί και ανταλλάξιμοι ως προς τη «φαινομενική συνείδηση» (phenomenal consciousness) αλλά όχι ως προς άλλες μορφές συνείδησης, όπως είναι η αυτοσυνείδηση.

- Εφόσον όμως χαιρείται για κάτι, το σκέφτεται, κι αυτό, έστω κι αν δεν υπάρχει, δεν είναι το ίδιο με το να μη σκέφτεται τίποτα.
- Το **περιεχόμενο λοιπόν της πεποιθήσής του υπάρχει σε ένα νοητικό χώρο** (στη συνείδησή του, ας πούμε) και **επηρεάζει τη συμπεριφορά του** (χαιρείται, πίνει το κώνειο, κλπ.).

Η αναγνώριση εσωτερικών νοητικών στοιχείων που μπορούν και επηρεάζουν τη συμπεριφορά αποτελεί κοινό στοιχείων πολλών ψυχολογικών θεωριών μετά τη δεκαετία του 1950. Προήλθε από την αντίδραση στις αδυναμίες της κυρίαρχης τότε συμπεριφοριστικής (μπαχαιοριστικής) ψυχολογίας, που αγνοούσε κάθε εσωτερικό παράγοντα στις απαντήσεις του οργανισμού και επικεντρωνόταν στην ανάλυση και τον έλεγχο των εξωτερικών ερεθισμάτων. Στην περίπτωση της αναπαραστατικής θεωρίας του νου τα εσωτερικά αυτά στοιχεία είναι οι γνώσεις, οι πεποιθήσεις, οι φόβοι, κλπ. των δρώντων υποκειμένων. Οι όροι αυτοί είναι δανεισμένοι από την «ψυχολογία του κοινού νου» ή «δημώδη ψυχολογία» (folk psychology), αν και, εκτός των γνώσεων, απαντώνται στην αναλυτική φιλοσοφία ως «προτασιακές ή αποβλεπτικές στάσεις» (propositional/intentional attitudes) (Brentano 1874/1973). Οι *νοητικές αναπαραστάσεις* είναι συμβολικές γλωσσικού-συντακτικού τύπου αναπαραστάσεις που έχουν ως περιεχόμενο τις γνώσεις ή τις προτασιακές στάσεις.

Η νοητική αναπαράσταση, είτε είναι αναπαράσταση για κάτι που υπάρχει είτε για κάτι που δεν υπάρχει, έχει αιτιολογικές ιδιότητες: στην περίπτωση του Σωκράτη προκαλεί χαρά (την ομολογεί ο ίδιος αλλά τη διαπιστώνουν και οι μαθητές του από το χαμόγελό του) και προκαλεί μια συμπεριφορά, πίνει το κώνειο. Η νοητική αναπαράσταση δηλαδή επιδρά στον υλικό κόσμο (στους μύες του προσώπου του όταν χαμογελάει και στους μύες του χεριού του όταν φέρνουν το κώνειο στο στόμα του) μέσω εγκεφαλικών περιοχών που ελέγχουν αυτές τις λειτουργίες.

Άρα η χαρά του Σωκράτη μπορεί να ερμηνευτεί με βάση την ενεργοποίηση του κέντρου ανταμοιβής του μεταχιαστικού συστήματος του εγκεφάλου (ή οποιουδήποτε άλλου κέντρου ελέγχει τη συγκινησιακή αντίδραση της χαράς) και η απόφασή του να πει το κώνειο από την ενεργοποίηση των μετωπιαίων προκινητικών περιοχών που προγραμματίζουν τη συμπεριφορά. Αυτή η ενεργοποίηση δεν μπορεί παρά να γίνει μέσω εισόδων από άλλες εγκεφαλικές περιοχές, όπου θα υποθέσουμε ότι *αναπαρίσταται* η προσμονή του για το επικείμενο ταξίδι του στα λημέρια των Θεών.

Αυτό σημαίνει ότι η οντολογία του νου δεν διαφέρει από την οντολογία του κόσμου: Οι νοητικές αναπαραστάσεις, εφόσον επηρεάζουν την ανθρώπινη συμπεριφορά, θα πρέπει να έχουν υλικό φορέα ώστε να επιδρούν *αιτιακά*. Η άποψη αυτή, κυρίαρχη σήμερα στη φιλοσοφία του νου και τη γνωσιακή επιστήμη, είναι γνωστή ως *θεωρία ταυτότητας*: Τα νοητικά φαινόμενα δεν είναι τίποτε άλλο παρά εγκεφαλικά φαινόμενα (αρκετά εύλογη υπόθεση εφόσον εγκεφαλικές βλάβες προκαλούν νοητικές διαταραχές) (Smart 1959, Feigl 1958, Place 1956).⁴

⁴ Η ταύτιση νοητικών καταστάσεων με εγκεφαλικές καταστάσεις μπορεί να είναι είτε *ταύτιση τύπου* (type-type identity) είτε *ταύτιση εκτύπου* (token-token identity). Στην πρώτη περίπτωση οι νοητικές καταστάσεις ταυτίζονται με νευροφυσιολογικές καταστάσεις, ενώ στη δεύτερη οι νοητικές καταστάσεις ταυτίζονται με λειτουργικές-υπολογιστικές καταστάσεις (κάθε νοητική κατάσταση ορίζεται με βάση τον αιτιακό της ρόλο, δηλαδή με βάση τον τρόπο που σχετίζεται με τις εισόδους, τις εξόδους και τις άλλες νοητικές καταστάσεις του οργανισμού). Η διαφορά ταύτισης τύπου και εκτύπου θα μπορούσε να φανεί με το παράδειγμα μιας σφραγίδας. Στην περίπτωση ταύτισης τύπου θα δεχόμασταν αναπαράσταση-αντίγραφο από μια και μόνη σφραγίδα, ενώ στην περίπτωση ταύτισης εκτύπου θα δεχόμασταν ως αναπαράσταση-αντίγραφο οποιοδή-

Οι νοητικές αναπαραστάσεις όμως, εκτός από τις *αιτιακές* ιδιότητες τους (επηρεάζουν τη συμπεριφορά) έχουν και *σημασιολογικές* ιδιότητες. Είναι δηλαδή αληθείς ή ψευδείς. Αν, για παράδειγμα, κάποιος πιστεύει ότι η Γη είναι επίπεδη, η αντίστοιχη νοητική αναπαράσταση αναφέρεται σε μια υπαρκτή πεποιθήση, η οποία όμως έχει ψευδές περιεχόμενο. Οι σημασιολογικές ιδιότητες των νοητικών αναπαραστάσεων θέτουν το πρώτο και σημαντικότερο πρόβλημα της αναπαραστασιακής θεωρίας του νου: Πώς μπορεί κάτι υλικό να έχει σημασιολογικό περιεχόμενο (να σημαίνει κάτι, να είναι αληθές ή ψευδές, καλό ή κακό, ωραίο ή άσχημο); (Fodor 1990).

Η απάντηση του Fodor είναι να μην ασχοληθούμε με το *ποιο* είναι το περιεχόμενο των αναπαραστάσεων (κατανοημένο ως συνθήκες αλήθειας) αλλά να καθορίσουμε το *τι* θεμελιώνει τη σχέση αναπαράστασης-αναπαριστώμενου. Με άλλα λόγια, το ερώτημα στο οποίο καλείται να απαντήσει η αναπαραστασιακή θεωρία είναι: δυνάμει τίνος πράγματος μια εσωτερική κατάσταση αποκτά αναπαραστασιακό περιεχόμενο (Fodor 1990, σελ. 318).

Ατομισμός ή ιντερναλισμός («Εφόσον ο εγκέφαλος είναι μέσα στο κεφάλι, ο νους δεν θα μπορούσε να βρίσκεται κάπου αλλού»).

Η απάντηση σε αυτό το πρόβλημα θα ήταν να θεωρήσουμε τη σχέση των νοητικών φαινομένων με τα εγκεφαλικά φαινόμενα ως *επιγένεση* (supervenience) των νοητικών φαινομένων από τα εγκεφαλικά, όπως οι αιτιακές ιδιότητες ενός οποιουδήποτε πράγματος επιγίνονται (supervene) στις φυσικές του ιδιότητες.⁵ Κατά συνέπεια, αν η ψυχολογία θέλει να είναι επιστημονική, θα πρέπει οι νοητικές καταστάσεις ενός ατόμου να επιγίνονται στις εγγενείς φυσικές του ιδιότητες: «Οι αιτιακές δυνάμεις επιγίνονται στις τοπικές μικροδομές. Στην ψυχολογική περίπτωση επιγίνονται στην τοπική νευρωνική δομή» (Fodor 1987, σελ. 45).

Σύμφωνα με τη θεωρία του ατομισμού, το περιεχόμενο των νοητικών αναπαραστάσεων μπορεί να περιγραφεί από τις αιτιακές σχέσεις τους, ανεξάρτητα από το αν υπάρχει ή όχι στο περιβάλλον του κάποιο συγκεκριμένο αντικείμενο ή ιδιότητα (ιντερναλισμός). Για την ατομιστική-ιντερναλιστική θεωρία του νοητικού περιεχομένου αυτό που έχει σημασία είναι το *πώς το υποκείμενο αναπαριστά τον κόσμο του και όχι πώς αλήθεια είναι ο κόσμος*. Αν πιστεύω, λόγου χάρη, ότι βλέπω ένα δέντρο και οι αιτιακές σχέσεις αυτής της νοητικής κατάστασης με τη συμπεριφορά με οδηγήσουν να συμπεριφερθώ σαν να βλέπω ένα δέντρο, τότε το περιεχόμενο της αντίστοιχης νοητικής αναπαράστασης είναι αληθές. Αυτό φαίνεται εύλογο, αν αναλογιστούμε ότι πολλές

ποτε όμοιο αποτύπωμα από παρόμοια σφραγίδα, μεταλλική, ξύλινη, πλαστική, με οποιαδήποτε χρώμα. Στην αντίστοιχη περίπτωση εγκεφάλου-αναπαράστασης, η θεωρία ταύτισης τύπου απαιτεί η αναπαράσταση να σχετίζεται με τη δομή των υποκείμενων εγκεφαλικών μηχανισμών, ενώ η θεωρία ταύτισης εκτύπου με τη λειτουργία τους. Στην τελευταία περίπτωση είναι δυνατή η ύπαρξη νοητικών καταστάσεων (ακόμη και συνειδησιακών) σε εγκεφάλους από τσιπ ή ηλεκτρονικούς υπολογιστές ή άλλες μηχανές που μιμούνται τη λειτουργία του εγκεφάλου

⁵ Ο όρος επιγένεση έχει χρησιμοποιηθεί στο παρελθόν από πολλούς ψυχολόγους και φιλοσόφους που διατύπωσαν την άποψη ότι η συνείδηση προκύπτει από την ύλη όταν αυτή φτάσει σε έναν υψηλό βαθμό συνθετότητας. Στη φιλοσοφία της νόησης εισήχθη γύρω στο 1970 από τον Davidson (1980): «...τα νοητικά χαρακτηριστικά είναι υπό μια έννοια εξαρτώμενα, ή επιγενόμενα, από τα φυσικά χαρακτηριστικά. Μια τέτοια επιγένεση θα μπορούσε να εκληφθεί ότι σημαίνει ότι δεν μπορούν να υπάρχουν δύο συμβάντα όμοια από κάθε φυσική άποψη αλλά διαφορετικά από κάποια νοητική άποψη, ή ότι ένα αντικείμενο δεν μπορεί να μεταβληθεί από κάποια νοητική άποψη χωρίς να μεταβληθεί από κάποια φυσική άποψη».

φορές έχουμε ψευδή συνείδηση, ψευδαισθήσεις ή πολλές φορές οι υπαρκτές νοητικές μας καταστάσεις αφορούν καταστάσεις πραγμάτων που δεν είναι υπαρκτές (π.χ. ο Δίας στην περίπτωση του Σωκράτη).

Με τον ατομισμό-ιντερναλισμό διατηρείται η φυσιοκρατική εξήγηση για το πώς αποκτούν οι νοητικές αναπαραστάσεις το περιεχόμενό τους. Στο βαθμό που οι νοητικές καταστάσεις προκαλούν αλλαγές στην παρατηρήσιμη συμπεριφορά και έχουν περιεχόμενο που επιγίνεται στις φυσικές ιδιότητες του εγκεφάλου, μπορούν να γίνουν αντικείμενο της νευροεπιστήμης. Εφόσον οι αναπαραστάσεις υπάρχουν κάπου στον εγκέφαλο και επηρεάζουν συγκεκριμένες δομές που σχετίζονται με τη συμπεριφορική έξοδο, θα μπορούσαμε ίσως στο μέλλον να τις εντοπίσουμε με εξελιγμένες νευροφυσιολογικές ή άλλες μεθόδους. Ένα όργανο λοιπόν των μελλοντικών νευροεπιστημόνων, το *εγκεφαλοσκόπιο*, θα μπορούσε να ανιχνεύσει τη νοητική αναπαράσταση από τις αιτιακές της επιδράσεις πάνω στη συμπεριφορά, θα εντόπιζε ανάδρομα τους εγκεφαλικούς μηχανισμούς που τη γεννούν και θα καθόριζε το περιεχόμενό της με βάση φυσικές ιδιότητες αυτών των μηχανισμών. Όσο φιλόδοξο και αν μοιάζει το πρόγραμμα ανάπτυξης του εγκεφαλοσκόπιου, είναι «θεωρητικά» δυνατό υπό τη θεωρία του ατομισμού και την παραδοχή της επιγένεσης.

Εξτερναλισμός («Οι σημασίες δεν βρίσκονται μέσα στο κεφάλι» -Putnam)

Τα πράγματα όμως περιπλέκονται αν ξαναγυρίσουμε στο ερώτημα αφετηρίας και ρωτήσουμε γιατί πραγματικά χαίρεται ο Σωκράτης. Ο Σωκράτης δεν χαίρεται γιατί απλώς θα πάει στα λημέρια του θεού του, αλλά γιατί θα πάει στα λημέρια του θεού του επειδή θα πει το κώνειο, υπακούοντας στην απόφαση του δικαστηρίου. Η υπακοή του στην απόφαση του δικαστηρίου είναι μια πράξη που «πρέπει» να γίνει και ο Σωκράτης χαίρεται επειδή υπακούει στην αίσθηση δικαίου που υπερασπίστηκε σε ολόκληρη τη ζωή του.

Οι μαθητές του όμως λυπούνται. Γι' αυτούς ο Σωκράτης «δεν πρέπει» να πει το κώνειο γιατί η απόφαση του δικαστηρίου ήταν άδικη. Το εγκεφαλοσκόπιό μας –ή μάλλον ο χειριστής του– θα μπερδεύεται. Θα ανίχνευε στο εγκέφαλο του Σωκράτη μια αναπαράσταση της πρότασης «πιστεύω ότι η υπακοή στην απόφαση του δικαστηρίου είναι δίκαιη» με αληθές περιεχόμενο (εφόσον οδηγεί σε συγκεκριμένη συμβατή συμπεριφορά –χαρά, την παραμονή στη φυλακή) και στον εγκέφαλο του Κρίτωνα μια αναπαράσταση της πρότασης «πιστεύω ότι η υπακοή στην απόφαση του δικαστηρίου είναι άδικη» με αληθές περιεχόμενο (εφόσον οδηγεί σε πάλι σε συμβατή συμπεριφορά –λύπη, προτροπή δραπέτευσης). Δυο λογικά αντίθετες προτάσεις, από τη σκοπιά του (εξωτερικού) τρίτου προσώπου, είναι και οι δύο αληθείς.

Ο υποστηρικτής της θεωρίας του ατομικισμού θα έλεγε ότι εφόσον αυτό που έχει σημασία είναι το πώς το υποκείμενο αναπαριστά τον κόσμο του και όχι πώς αλήθεια είναι ο κόσμος, η έννοια της ψευδούς συνείδησης αποτελεί πρόβλημα του εγκεφαλοσκόπιου ή του χρήστη του και όχι των δρώντων υποκειμένων.

Εάν όμως το σύμβολο δεν παραπέμπει από μόνο του στο αναπαριστώμενο –εάν δηλαδή μια νοητική αναπαράσταση δεν παραπέμπει από μόνη της στην αλήθεια της πρότασης που έχει ως περιεχόμενό της– και χρειάζεται κάποιον ερμηνευτή-χρήστη, τότε χάνεται η φυσιοκρατική εξήγηση. Σε αυτήν την περίπτωση θα χρησιμοποιούνταν δηλαδή οι προτασιακές στάσεις του χρήστη-ερμηνευτή για να εξηγηθεί η λειτουργία και το περιεχόμενο των νοητικών αναπαραστάσεων. Για να λυθούν τα παράδοξα των ψευδών

αναπαραστάσεων, των διαφορετικών αναπαραστάσεων από παρόμοια ερεθίσματα και η ανάγκη του χρήστη-ερμηνευτή (για να διατηρηθεί δηλαδή η φυσικαλιστική ερμηνεία του περιεχομένου των αναπαραστάσεων) η αναπαραστατική θεωρία του νου φαίνεται να ευνοεί τις αποκαλούμενες *εξτερναλιστικές προσεγγίσεις*.

Σύμφωνα με τους υποστηρικτές του εξτερναλισμού, το περιεχόμενο των νοητικών αναπαραστάσεων είναι «ευρύ», δεν εξαρτάται μόνο από τα εγγενή (intrinsic) χαρακτηριστικά του υποκειμένου, αλλά και από την ιστορία του και το φυσικό και κοινωνικό του περιβάλλον. Αφετηρία των εξτερναλιστικών απόψεων αποτελεί το περιβόητο νοητικό πείραμα της «δίδυμης Γης» του Hillary Putnam (1975),⁶ με το οποίο έδειξε ότι η σημασία των λέξεων δεν καθορίζεται από τις ψυχολογικές καταστάσεις του χρήστη των λέξεων: «meanings just ain't in the head» (Putnam 1975, σελ. 227).

Στο πλαίσιο της αναπαραστατικής θεωρίας η σημασιολογία της γλώσσας είναι εύκολο να μετατραπεί σε σημασιολογία της νόησης, επειδή σύμφωνα με την επικρατούσα φοντοριανή νοητική αρχιτεκτονική ο νους είναι δομημένος ως γλώσσα. Ότι ισχύει λοιπόν για τις λέξεις της φυσικής γλώσσας ισχύει και για τις «λέξεις» της νοητικής γλώσσας, δηλαδή τις νοητικές αναπαραστάσεις. Από αυτήν την αναλογία προέκυψε το σύνθημα της εξτερναλιστικής θεωρίας: «οι νοητικές αναπαραστάσεις δεν είναι μέσα στο κεφάλι» (McGinn 1989, σελ. 31).

Σύμφωνα λοιπόν με την αντι-ατομιστική θεωρία του εξτερναλισμού, το υποθετικό μας εγκεφαλοσκόπιο δεν μπορεί να αποκαλύψει το ευρύ περιεχόμενο (τις σημασίες) των νοητικών αναπαραστάσεων, γιατί αυτό δεν βρίσκεται μέσα στο κεφάλι. Στην περίπτωση των πανομοιότυπων ανθρώπων που ζουν στη Γη και τη δίδυμη Γη του Putnam, δεν μπορεί από τις σχετικές αναπαραστάσεις του νερού μέσα στον εγκέφαλό τους να αποδώσει το σωστό περιεχόμενο H₂O ή XYZ εάν δεν διαθέτει μια βάση δεδομένων εξωτερική ως προς τα δύο υποκείμενα, γιατί η αλήθεια (η σημασία) αυτού του περιεχομένου βρίσκεται στο *φυσικό περιβάλλον*. Στην περίπτωση του Σωκράτη και των μαθητών του, το εγκεφαλοσκόπιο δεν μπορεί να ανιχνεύσει τη σημασία (την αλήθεια) του περιεχομένου των πεποιθήσεων που προκαλούν τη χαρά και λύπη αντίστοιχα, γιατί αυτή βρίσκεται σε έννοιες και αξίες που αναπαρίστανται μεν στον εγκέφαλό τους, αλλά προέρχονται από το *κοινωνικό περιβάλλον*.

⁶ Ο Α και ο Β είναι δυο πανομοιότυποι άνθρωποι μόριο προς μόριο, που βρίσκονται ο ένας στη Γη και ο άλλος σε έναν πλανήτη ενός «δυνατού» κόσμου –τη δίδυμη Γη- ο οποίος είναι φαινομενικά πανομοιότυπος με τη Γη. Επίσης, τόσο ο Α όσο και ο Β, οι οποίοι δεν γνωρίζουν τη χημική σύσταση του νερού, όταν αναφέρονται στο νερό χρησιμοποιούν τη λέξη «νερό». Ωστόσο, οι δύο πλανήτες, παρότι φαινομενικά (δηλαδή μακροσκοπικά) πανομοιότυποι, διαφέρουν ως προς τη μοριακή σύσταση του νερού: στη δίδυμη Γη το νερό δεν έχει τη χημική σύσταση H₂O αλλά XYZ. Παρότι, δηλαδή, το νερό έχει τις ίδιες μακροσκοπικές ιδιότητες –είναι ένα άχρωμο άοσμο υγρό που ξεδιψά, σβήνει τη φωτιά κλπ.- και προσδιορίζεται με την ίδια περιγραφή, η πραγματική του υπόσταση διαφέρει στους δύο πλανήτες. Άρα η αναφορά (το περιεχόμενο) της λέξης «νερό» όταν εκφέρεται από τον Β δεν είναι το H₂O αλλά το XYZ. Έτσι λοιπόν, ενώ ο Α και ο Β βρίσκονται σε πανομοιότυπες ψυχολογικές καταστάσεις, όταν προφέρουν τη λέξη «νερό» αναφέρονται σε διαφορετικά πράγματα.

Η αναγωγή της συνείδησης στις αναπαραστάσεις και το νευρικό τους υπόστρωμα⁷

Σύμφωνα με μια ευρέως αποδεκτή θεωρία του Ned Block (1995), η συνείδηση διακρίνεται σε *φαινομενική συνείδηση* (phenomenal consciousness), *προσβαίνουσα συνείδηση* ή *συνείδηση πρόσβασης* (access c.), *αυτοσυνείδηση* (self-consciousness) και *ελέγχουσα ή εσωτερική συνείδηση* (monitor c.).

Η φαινομενική συνείδηση ταυτίζεται με την εμπειρία και οι αντίστοιχες συνειδητές καταστάσεις είναι καταστάσεις με εμπειρικές ιδιότητες. Το σύνολο των ιδιοτήτων αυτών μιας κατάστασης ισοδυναμεί με το «*πώς είναι να βλέπεις ένα χρώμα, να ακούς έναν ήχο, να έχεις μια γεύση, να αισθάνεσαι πόνο, χαρά, κλπ.*». Το περιεχόμενο της φαινομενικής συνείδησης ταυτίζεται με τις ποιότητες (qualia).

Μια κατάσταση ανήκει στην προσβαίνουσα συνείδηση εάν το υποκείμενο διαθέτει μια αναπαράσταση του περιεχομένου της που (α) χρησιμοποιείται ως προκείμενη σε έναν συλλογισμό, (β) χρησιμοποιείται στο λογικό έλεγχο μιας πράξης και (γ) χρησιμοποιείται στο λογικό έλεγχο του λόγου. Με αυτήν την έννοια η προσβαίνουσα συνείδηση περιλαμβάνει τα συστήματα μνήμης, την αντιληπτική κατηγοριοποίηση, τη λήψη αποφάσεων, το σχεδιασμό, την εκτίμηση εναλλακτικών λύσεων, την εκούσια επικέντρωση της προσοχής και γενικότερα τον ορθολογικό έλεγχο της συμπεριφοράς.

Η αυτοσυνείδηση είναι «η κατοχή μιας έννοιας του εαυτού και η δυνατότητα χρήσης αυτής της έννοιας στις σκέψεις για τον εαυτό μας» (Block 1995, σελ. 235). Η ελέγχουσα (εσωτερική) συνείδηση αντιστοιχεί σε τρεις έννοιες: την εσωτερική αντίληψη, την εσωτερική διερεύνηση και την αποκαλούμενη «σκέψη υψηλής τάξης».

⁷ Αν και το θεωρητικό υπόβαθρο της αναγωγής έχει συζητηθεί με αρκετή σαφήνεια από τον Α. Μοσχοβάκη στον πρώτο κύκλο, θα ήθελα να προσθέσω μια σύντομη αναφορά που αφορά ειδικότερα το θέμα της αναγωγής της αναπαραστατικής θεωρίας του νου στις νευροεπιστήμες. Σύμφωνα με την κλασική περιγραφή του Ernest Nagel (1961) η **διαθεωρητική αναγωγή** είναι πλήρης όταν οι νόμοι μιας θεωρίας προκύπτουν λογικά από τους νόμους μιας άλλης με τις κατάλληλες «αρχές γεφύρωσης» που συνδέουν τις εκφράσεις της μιας με τις εκφράσεις της άλλης. Στην περίπτωση της αναγωγής της ψυχολογίας στη νευροεπιστήμη, εάν η μία θεωρία είναι η αναπαραστατική θεωρία της νόησης και η άλλη η νευροφυσιολογία, οι αρχές γεφύρωσης θα πρέπει να συνδέουν τις νοητικές ιδιότητες με τις νευροφυσιολογικές ιδιότητες. Εάν η σύνδεση είναι διπλά-εξαρτώμενη (biconditional), τότε πρόκειται για **πλήρη αναγωγή**. Εάν η σύνδεση συνιστά μονοσήμαντη συσχέτιση, κατά τη οποία οι ιδιότητες του ανωτέρου επιπέδου εξαρτώνται οντολογικά και προσδιορίζονται από τις ιδιότητες του κατώτερου επιπέδου χωρίς να συμβαίνει το αντίθετο, τότε πρόκειται για **επιγένεση** και όχι για πλήρη αναγωγή. Λόγω της οντολογικής σχέσης των ιδιοτήτων του ανώτερου επιπέδου με αυτές του κατώτερου, επειδή οι τελευταίες είναι υλικής φύσης, τότε και οι επιγενόμενες έχουν υλιστικό χαρακτήρα. Η επιγένεση δηλαδή είναι μια μορφή μη πλήρως αναγωγικού φυσικαλισμού. Μία άλλη μορφή μη αναγωγικού φυσικαλισμού είναι και η **ανάδυση** (emergentism), κατά την οποία οι ιδιότητες του ανωτέρου επιπέδου αποτελούν νέες «γνήσιες» ιδιότητες που δεν εξαρτώνται από αυτές του κατώτερου επιπέδου. Όπως θα φανεί στη συνέχεια, η οντολογική αρχή που ακολουθούμε σε αυτήν την παρουσίαση είναι ο «**ανώμαλος μονισμός**» του Donald Davidson (1980), σύμφωνα με τον οποίον η ταυτότητα νοητικών και εγκεφαλικών γεγονότων είναι ταυτότητα εκτύπων (συγκεκριμένων νοητικών και φυσικών συμβάντων) και όχι τύπων (γενικού είδους γεγονότων). (Ο ανώμαλος μονισμός είναι σχεδόν μονόδρομος για την αναπαραστατική θεωρία του νου λόγω του λειτουργιστικού της χαρακτήρα και του βάρους της επίδρασης του Fodor στη διατύπωσή της.) Σύμφωνα με τον Davidson, όταν έχω την εμπειρία ενός πόνου, ο πόνος αυτός ταυτίζεται με κάποια συγκεκριμένη φυσική κατάσταση μέσα μου, ενώ όταν εσύ πονάς, ο πόνος σου ταυτίζεται με κάποια συγκεκριμένη φυσική κατάσταση μέσα σου, χωρίς αυτό να σημαίνει ότι όταν πονάμε και οι δύο μοιραζόμαστε μια κοινή φυσική κατάσταση. Ο ανώμαλος μονισμός είναι μια μη εξαιρετική θεωρία ταυτότητας, η οποία όμως δεν πληροί τα κριτήρια του Nagel ως πλήρης αναγωγική θεωρία.

Ας επανέλθουμε τώρα στο ερώτημα αφετηρίας και να ελέγξουμε εάν με τη βοήθεια της έννοιας της αναπαράστασης και των νευρωνικών της ισοδυνάμων μπορούμε να απαντήσουμε στο ερώτημα πώς ο εγκέφαλος «παράγει» τις παραπάνω συνειδητές καταστάσεις. Κατ' αρχάς, οι περισσότερες συνειδητές καταστάσεις έχουν κάποιο είδος αναπαραστασιακού περιεχομένου, αναπαριστώντας τον κόσμο με τον έναν ή τον άλλον τρόπο. Η ιδιαίτερη φύση της οπτικής συνειδητής εμπειρίας, για παράδειγμα, μπορεί να χαρακτηριστεί με όρους περιεχομένου. Μια οπτική εμπειρία αναπαριστά τον κόσμο *σαν να*⁸ περιέχει σχήματα και χρώματα, *σαν να* περιέχει διάφορα αντικείμενα που βρίσκονται σε διάφορες σχέσεις, κ.ο.κ. Εάν η αντίληψη είναι αληθινή, τότε ο κόσμος είναι όπως τον αναπαριστά η εμπειρία. Εάν πρόκειται για ψευδαίσθηση, τότε ο κόσμος δεν είναι όπως τον αναπαριστά η εμπειρία. Και στις δυο περιπτώσεις, όμως, οι οπτικές εμπειρίες έχουν σαφές αναπαραστασιακό περιεχόμενο.⁹ Το ίδιο ισχύει και για τις υπόλοιπες αισθητικές εμπειρίες, καθώς και για πολλές ή τις περισσότερες μη αισθητικές εμπειρίες.

Όσον αφορά στο περιεχόμενο των αισθητικών εμπειριών (qualia) υπάρχει μια πληθώρα ερευνών και ενδείξεων ότι μπορούν να συσχετιστούν με τη νευρωνική δραστηριότητα συγκεκριμένων εγκεφαλικών περιοχών. Τα πειράματα του Νίκου Λογοθέτη με τον διοφθάλμιο ανταγωνισμό σε πιθήκους (Logothetis and Schall 1986, Leopold and Logothetis 1996, Steinberg and Logothetis 1997) έδειξαν ότι στην κάτω κροταφική περιοχή (inferior temporal area - IT) υπάρχουν κύτταρα που εκφορτίζονται όταν οι πίθηκοι συμπεριφέρονται «σαν να βλέπουν» σχήματα, ανεξάρτητα από την περιοχή V1, τα κύτταρα της οποίας σχετίζονται με την είσοδο από τον αμφιβληστροειδή και όχι με την οπτική αντίληψη. Επίσης, κύτταρα στη μέση κροταφική περιοχή (MT/V5) σχετίζονται με την αντίληψη της κίνησης (Heeger et al. 1999), καθώς και με την αντίληψη κινητικών μεταειδώλων, που στην ουσία παριστούν ψευδή αντίληψη (Huk et al. 2001). Αμφοτερόπλευρες βλάβες στην MT/V5 προκαλούν ακινητοψία, αδυναμία δηλαδή αντίληψης της κίνησης (Rees et al. 2002).

Έχει προταθεί (Block 2005, Rees et al. 2002) ότι οι δύο παραπάνω περιοχές αποτελούν το «ελάχιστο νευρωνικό αντίστοιχο της συνείδησης» (ENΑΣ) σε ό,τι αφορά την οπτική εμπειρία. Οι δύο συγκεκριμένες περιοχές κωδικοποιούν αναπαραστάσεις που το περιεχόμενό τους σχετίζεται (correlates) με το αναπαραστασιακό περιεχόμενο της φαινομενικής συνειδητής οπτικής εμπειρίας. Οι παραπάνω ερευνητές προσφέρουν ενδείξεις ότι οι προβολές των περιοχών αυτών σε συγκεκριμένες περιοχές του μετωπιαίου και του βρεγματικού λοβού παριστούν το νευρωνικό αντίστοιχο της προσβαίνουσας συνείδησης, που σχετίζεται με το σχεδιασμό και τον έλεγχο της συνειδητής συμπεριφοράς.¹⁰

⁸ Οι φιλόσοφοι χρησιμοποιούν την έκφραση «σαν να» επειδή η εμπειρία μπορεί να συμβαίνει –και συμβαίνει– ακόμη κι όταν δεν υπάρχουν τα αντίστοιχα εξωτερικά ερεθίσματα.

⁹ Αυτό αποτελεί και το πλεονέκτημα της αναπαραστασιακής θεωρίας της συνείδησης (βλ. Lycan 2005), επειδή συχνά στη συζήτηση για τα qualia αναφέρεται η περίπτωση ψευδών αντιλήψεων, όπως π.χ. χρωματικών μεταειδώλων χωρίς τα αντίστοιχα εξωτερικά ερεθίσματα, που για ορισμένους φιλοσόφους αποτελεί εμπόδιο στις φυσικαλιστικές ερμηνείες.

¹⁰ Εκτός των παραπάνω πρόσφατων ερευνών, υπάρχουν αρκετές προτάσεις για το «ελάχιστο νευρωνικό αντίστοιχο της συνείδησης» (ENΑΣ) ως το «ελάχιστο σύνολο νευρωνικών γεγονότων και μηχανισμών επαρκών για ένα συγκεκριμένο συνειδητό αντίλημα», όπως το έχει ορίσει ο Koch (2004, σελ. 16). Για παράδειγμα, ο Libet (1993) προτείνει ως ENΑΣ τη χρονικά παρατεταμένη νευρωνική πυροδότηση (temporally extended neural firing). Οι Crick and Koch (1998) θεωρούν ότι το ENΑΣ μπορεί να περιλαμβάνει ένα ιδιαίτερο είδος κυττάρων στο σύνολο του φλοιού. Ο Edelman (1989) προτείνει ότι το ENΑΣ μπορεί να

Πώς όμως αυτά τα αναπαραγωγίμα πειράματα σε πιθήκους ή ανθρώπους μπορούν να απαντήσουν στο θεμελιώδες ερώτημα της υποκειμενικότητας των εμπειριών (qualia) και να αποτελέσουν πραγματικά έναν επιστημονικό τρόπο μελέτης της συνείδησης; Η αναμφισβήτητα υποκειμενική φύση των εμπειριών έχει οδηγήσει ορισμένους συγγραφείς να παραδεχτούν ένα «ερμηνευτικό χάσμα» (explanatory gap) ανάμεσα στην επιστημονική περιγραφή και την υποκειμενική εμπειρία (Block and Stalnaker 1999, Levine 1983). Άλλοι το θεωρούν ως το «δύσκολο πρόβλημα» της συνείδησης, διαχωρίζοντάς το από το «εύκολο πρόβλημα» της περιγραφής της λειτουργίας του εγκεφάλου σε συνειδητές ή μη συνειδητές καταστάσεις (Chalmers 1995). Ο Chalmers (1995) έχει χρησιμοποιήσει το νοητικό πείραμα με τη Μαίρη για να δείξει ότι το «δύσκολο πρόβλημα» της συνείδησης αφορά ακριβώς αυτή τη διάκριση υποκειμενικής εμπειρίας-αντικειμενικών εγκεφαλικών μηχανισμών (πώς οι φυσικές διαδικασίες στον εγκέφαλο παράγουν την υποκειμενική εμπειρία). Υποθέτει ότι η Μαίρη είναι μία νευροεπιστήμονας του 23^{ου} αιώνα, η οποία γνωρίζει τα πάντα για τους οπτικούς μηχανισμούς του εγκεφάλου, αλλά ζει σε ένα λευκό δωμάτιο χωρίς κανένα χρώμα. Γνωρίζει πώς ο εγκέφαλος διακρίνει τα χρώματα, γνωρίζει τα ονόματα των χρωμάτων που αντιστοιχούν σε συγκεκριμένα μήκη κύματος, γνωρίζει πώς ο εγκέφαλος παράγει τη λεκτική τους αναφορά. Δεν γνωρίζει όμως πώς είναι να έχεις την εμπειρία ενός χρώματος, π.χ. του κόκκινου. Κατά συνέπεια, συμπεραίνει ο Chalmers, η συνειδητή εμπειρία δεν ανάγεται στα φυσικά γεγονότα της λειτουργίας του εγκεφάλου. Στην απάντησή τους, όμως, οι Crick and Koch (1995) αναφέρουν ότι η Μαίρη δεν έχει την εμπειρία του κόκκινου επειδή, απλώς, *ποτέ δεν είχε μια νευρωνική αναπαράσταση του κόκκινου χρώματος*. Αλλά και αν είχε μια αναπαράσταση του κόκκινου, η εμπειρία της αυτή θα ήταν *υποκειμενική*, με την έννοια ότι δεν θα μπορούσε να εκφραστεί με λόγια. Η νευρωνική πληροφορία από την αναπαράσταση μεταφέρεται στις κινητικές περιοχές του εγκεφάλου, όπου γίνεται διαθέσιμη για την πράξη ή για το λόγο. Κατά τη μεταφορά αυτή πάντοτε γίνεται επανακωδικοποίηση των πληροφοριών, έτσι ώστε η πληροφορία που εκφράζεται από τους κινητικούς νευρώνες σχετίζεται αλλά δεν ταυτίζεται με την πληροφορία που εκφράζεται από τους νευρώνες που εκφράζουν την οπτική εμπειρία, σε κάποιο επίπεδο της οπτικής ιεραρχίας. Μια πλευρά της υποκειμενικότητας λοιπόν συνίσταται σε αυτήν την αδυναμία να μεταφερθεί με λέξεις και ιδέες το περιεχόμενο της συνειδητής εμπειρίας (λόγω της πληροφοριακής επανακωδικοποίησης) και όχι σε κάποια οντολογική διάκριση υποκειμένου της εμπειρίας και αντικειμένου της επιστημονικής περιγραφής.

Είναι λοιπόν επιστημονικό το ερώτημα εάν το κόκκινο που βλέπω εγώ είναι *υποκειμενικά* το ίδιο με το κόκκινο που βλέπεις εσύ (έχουν δηλαδή την ίδια ποιότητα-qualia: κοκκινότητα); Εάν το νευρωνικό αντίστοιχο της αναπαράστασης του κόκκινου είναι ακριβώς το ίδιο στο δικό σου εγκέφαλο και το δικό μου εγκέφαλο, τότε είναι επιστημονικά λογικό να συμπεράνω ότι βλέπεις το κόκκινο όπως το βλέπω εγώ. Το πρόβλημα βρίσκεται στη λέξη «ακριβώς». Το πόσο ακριβείς (πρέπει να) είμαστε, εξαρτάται από τη λεπτομερή γνώση των διαδικασιών που λαμβάνουν μέρος. Εάν το νευρωνικό αντίστοιχο του κόκκινου εξαρτάται, σε σημαντικές παραμέτρους, από την προηγούμενη εμπειρία μου, και εάν η προηγούμενη εμπειρία μου είναι σημαντικά διαφορετική από τη δική σας, τότε μπορεί να μην είμαστε σε θέση να συμπεράνουμε ότι

περιλαμβάνει αμφίδρομους θαλαμοφλοιώδεις βρόχους. (Βλ. Ε. Σαββάκη εισήγηση στον πρώτο κύκλο και Chalmers (1998, 2000) για έναν πλήρη κατάλογο των σχετικών ερευνών και προτάσεων).

βλέπουμε το κόκκινο με τον ίδιο ακριβώς τρόπο (Crick 1994, σελ. 31). Η υποκειμενική διάσταση δεν αφορά την (μυστική ή απροσπέλαστη) φύση του περιεχομένου της εμπειρίας αλλά το «πληροφοριακό» της εύρος.

Εάν λοιπόν το πώς βλέπω το κόκκινο εξαρτάται από παρελθούσες εμπειρίες μου, αλλά και γνώσεις, στάσεις ή πεποιθήσεις γι' αυτό που βλέπω, όλες αυτές οι ψυχολογικές καταστάσεις αναπαρίστανται στον εγκέφαλό μου και οι αναπαραστάσεις αυτές έχουν ως περιεχόμενο αντίστοιχες μνημονικές εγγραφές. Το περιεχόμενο (η ποιότητα, το *quale* κοκκινότητα) της τρέχουσας εμπειρίας μου είναι το αποτέλεσμα του «συνδυασμού» όλων αυτών των αναπαραστάσεων και είναι φυσικό να είναι διαφορετική από τη δική σας, επειδή οι αναπαραστάσεις μου αυτές είναι διαφορετικές από τις δικές σας. Δεν υπάρχει τίποτα το μη φυσικό, μη υλικό ή μη επιστημονικό σε αυτή τη διαδικασία. Ούτε είναι ανάγκη να καταφύγουμε σε μεγάλους αριθμούς, σε στοχαστικές ή χαοτικές διαδικασίες (βλ. Α. Μοσχοβάκης, εισήγηση στον πρώτο κύκλο) για να ερμηνεύσουμε την υποκειμενικότητα. Ακόμη και μία προηγούμενη εμπειρία μου να είναι διαφορετική από τη δική σας, η τρέχουσα θα είναι και αυτή διαφορετική.

Η άποψη αυτή, ότι δηλαδή η υποκειμενικότητα οφείλεται σε σχέσεις και συνδυασμούς διαφορετικών αναπαραστάσεων, είναι απόλυτα συμβατή με την ιντερναλιστική εκδοχή της αναπαραστατικής θεωρίας του νου, κατά την οποία το (στενό) περιεχόμενο μιας αναπαράστασης εξαρτάται από τις σχέσεις της με άλλες αναπαραστάσεις και τη συμπεριφορά. Στο χώρο της εμπειρικής έρευνας, ωστόσο, τίθεται το ερώτημα του πώς υλοποιούνται φυσικά αυτοί οι συνδυασμοί. Πώς τα υπόλοιπα μέρη του εγκεφάλου (π.χ. εκεί που αναπαρίστανται προηγούμενες εμπειρίες) γνωρίζουν ότι η πυροδότηση ενός συγκεκριμένου συνόλου νευρώνων παράγει το συνειδητό αντίλημμα (percept) ενός χρώματος ή ενός προσώπου; Πώς γνωρίζει ο εγκέφαλος τι αναπαριστά η πυροδότηση αυτών των νευρώνων, ποιο είναι το περιεχόμενο της αναπαράστασης;

Η απάντηση πολλών νευροεπιστημόνων (βλ. Crick and Koch 1998, σελ. 104) είναι ότι το περιεχόμενο προκύπτει ακριβώς από τη συγχρονισμένη πυροδότηση των συγκεκριμένων νευρώνων και τη σύνδεση με άλλες σχετικές αναπαραστάσεις, καθώς και με την έξοδο-συμπεριφορά. Για παράδειγμα, οι νευρώνες που σχετίζονται με την οπτική αντίληψη κάποιου προσώπου μπορεί να συνδέονται με άλλους που αναπαριστούν το όνομά του, με άλλους που αναπαριστούν τη φωνή του, με μνήμες σχετικά με αυτόν ή αυτήν, σε ένα ευρύ συνειρμικό δίκτυο, παρόμοιο με μια σχεσιακή βάση δεδομένων.¹¹ Το τι σημαίνει μια κόκκινη ντομάτα ή το γνωστό ή άγνωστο πρόσωπο που βλέπω εξαρτάται από το ποιες και πόσες αναπαραστάσεις ενεργοποιούνται από την οπτική αντίληψη (γι' αυτό και το *binding* αποτελεί σημαντική ιδιότητα –βλ. Llinas [2001, σελ 126] “It binds, therefore I am”).¹²

¹¹ Το πρόβλημα του πώς σχετίζονται και συνδέονται οι διαφορετικές αυτές αναπαραστάσεις είναι γνωστό ως «πρόβλημα σύνδεσης» (*binding problem*) (Roskin 1999). Μία αρκετά δημοφιλής λύση του προβλήματος της σύνδεσης βασίζεται στην ανακάλυψη ότι γειτονικοί ή πιο απομακρυσμένοι πληθυσμοί νευρώνων τείνουν να συγχρονίζουν την πυροδότησή τους με μια περίοδο συγχρονισμού από 35 έως 75 Hz (Gray and Singer 1989, Eckhorn et al. 1988). Αυτός ο συγχρονισμός, που έχει ονομαστεί «ταλαντώσεις γ» (*gamma-oscillations*), έχει θεωρηθεί τόσο σημαντικός ώστε έχει προταθεί ο ίδιος ως το νευρωνικό αντίστοιχο της συνείδησης (Crick and Koch 1990).

¹² Εάν η αισθητική είσοδος οδηγείται στην έξοδο χωρίς ενεργοποίηση άλλων αναπαραστάσεων, τότε μιλάμε για αισθητικοκινητικό αντανακλαστικό. Όσο οι εγκέφαλοι γίνονται πιο πολύπλοκοι τόσο οι αναπαραστασιακές σχέσεις γίνονται πιο πλούσιες και τα περιεχόμενα αποκτούν «χρoιές» (υποκειμενική «χρoιά» στον άνθρωπο).

Είναι λοιπόν σαφές και πάλι ότι όλες οι νευροεπιστημονικές εξηγήσεις εντάσσονται στο θεωρητικό πλαίσιο της ατομιστικής-ιντερναλιστικής θεωρίας, κατά την οποία το περιεχόμενο των αναπαραστάσεων (και κατ' επέκταση των συνειδητών καταστάσεων) προκύπτει από τις σχέσεις τους με άλλες αναπαραστάσεις. Όπως είδαμε όμως με το πείραμα της δίδυμης Γης, η ίδια η έννοια του αναπαραστασιακού περιεχομένου επιτρέπει την ιδέα των διπλών καταστάσεων (duplicate states) – καταστάσεων δηλαδή με το ίδιο περιεχόμενο που διαφέρουν ως προς τη σημασία τους (με όρους τιμής αλήθειας). Το τι αναπαριστά μια αναπαράσταση είναι ένα πράγμα, το αν είναι αληθής άλλο. Π.χ. μια αναπαράσταση μπορεί να αναπαριστά ένα αντικείμενο που υπάρχει ή ένα αντικείμενο που δεν υπάρχει. Το νευρωνικό αντίστοιχο και στις δυο περιπτώσεις θα είναι το ίδιο και η σημασία (η αλήθεια) της αναπαράστασης ελέγχεται μόνο από έναν εξωτερικό παρατηρητή (π.χ. τον πειραματιστή στα πειράματα με τον διοφθάλμιο ανταγωνισμό που αναφέραμε παραπάνω).

Η (νευρο)επιστημονική λοιπόν μελέτη των συνειδητών καταστάσεων αφορά τη «στενή» εκδοχή του περιεχομένου που εξαντλείται στις αιτιακές σχέσεις με άλλες αναπαραστάσεις και δεν μπορεί να περιλάβει το ευρύ περιεχόμενο και τις σημασιολογικές του σχέσεις. Η νευροφυσιολογία μπορεί να απαντήσει (τώρα ή στο μέλλον, με νευροβιολογικά ή υπολογιστικά μοντέλα) στο ερώτημα «πώς παράγονται οι υποκειμενικές συνειδητές εμπειρίες μέσα στον εγκέφαλο» από την πυροδότηση των κυττάρων του νευρωνικού τους υποστρώματος, κατανέμοντας το πρόβλημα αυτό σε επιμέρους ερωτήματα, όπως π.χ. πώς παράγεται ένα συνειδητό αντίλημα (percept), πώς συνδυάζεται (binds) με άλλες αναπαραστάσεις ώστε να αποκτήσει φαινομενικό συνειδητό περιεχόμενο, πώς επανακωδικοποιείται στην πορεία του προς τα συστήματα εξόδου.

Μέσα σε αυτήν την αιτιακή ακολουθία γεγονότων το περιεχόμενο (η σημασία) των συνειδητών καταστάσεων εξαντλείται πράγματι στις αιτιακές σχέσεις. Εάν έχω την πεποίθηση ότι η ντομάτα που βλέπω είναι κόκκινη ή ότι η Γη είναι επίπεδη, οι αναπαραστάσεις που έχουν αυτές τις πεποιθήσεις ως περιεχόμενο κωδικοποιούνται νευρικά, συνδυάζονται με άλλες αναπαραστάσεις και χρησιμοποιούνται σε συλλογισμούς, στη συμπεριφορά ή στη γλώσσα (στην προσβαίνουσα συνείδηση). Καμία όμως νευροεπιστημονική τεχνική δεν μπορεί να αποφανθεί μόνη της, χωρίς την κρίση του παρατηρητή ή του πειραματιστή, αν είναι αληθής ή ψευδής ως προς το ευρύ της περιεχόμενο, ως προς τη σχέση της με τον πραγματικό κόσμο. Είτε αληθής είτε ψευδής, το νευρωνικό υπόστρωμα είναι το ίδιο. Το ερμηνευτικό χάσμα πρέπει να αποδοθεί σε αυτήν ακριβώς την επιστημολογική διάκριση στενού και ευρέως περιεχομένου και όχι σε κάποια οντολογική διάκριση της υποκειμενικότητας των εμπειριών και της αντικειμενικότητας των υποκειμένων μηχανισμών.

Η ευρεία σημασία δεν «παράγεται» στον εγκέφαλο. Κατ' επέκταση, το ερώτημα «πώς κάτι υλικό μπορεί να έχει σημασιολογικές ιδιότητες» δεν απαντάται από τη νευροεπιστήμη. Οι εξτερναλιστικές εκδοχές της αναπαραστατικής θεωρίας μπορεί να δίνουν κάποιες ερμηνείες (π.χ. η αιτιακή-ιστορική θεωρία της αναφοράς του Putnam [1975], η πληροφοριακή σημασιολογία του Dretske [1981], η θεωρία της ασύμμετρης εξάρτησης του Fodor [1990] και η τελεολογική σημασιολογία της Millikan [1984, 1993]), αυτές όμως δεν μπορούν να ελεγχθούν νευροφυσιολογικά. Στην καλύτερη περίπτωση μπορούν να γίνουν αντικείμενο διεπιστημονικής προσέγγισης (βλ. Sutton 2004). Και σ' αυτήν την περίπτωση όμως υπάρχουν περιορισμοί. Υπάρχουν

αναπαραστάσεις που έχουν μόνο ευρύ περιεχόμενο, περιέχοντας σχέσεις του υποκειμένου με το περιβάλλον του, που μπορεί να είναι γνώσεις του είδους (με τη δαρβινική έννοια) ή γνώσεις του ατόμου, μαθημένες συμπεριφορές ή κρίσεις, που δεν έχουν την κλασική συμβολική μορφή αναπαραστάσης-αναπαριστώμενου.¹³ Οι αναπαραστάσεις αυτές δεν αφορούν άμεσες αισθητικές εμπειρίες, δεν ανήκουν δηλαδή στη φαινομενική συνείδηση, αλλά στην προσβαίνουσα, επειδή αφορούν συνειδητές γνώσεις, καταστάσεις και στάσεις που (υπερ)καθορίζουν και ελέγχουν τη συμπεριφορά.¹⁴ Ιδιαίτερα οι αναπαραστάσεις πεποιθήσεων που βασίζονται σε αξιολογικές κρίσεις (π.χ. [αυτό] είναι καλό, [αυτό] είναι δίκαιο, [αυτό] είναι ωραίο) έχουν ένα σημασιολογικό περιεχόμενο που δεν μπορεί να διερευνηθεί με όρους αλήθειας, άρα δεν μπορεί να γίνει αντικείμενο επιστημονικής έρευνας. Εάν λ.χ. ο Σωκράτης είχε την πεποίθηση ότι η Γη είναι επίπεδη, αυτή είναι *αληθής ως προς το στενό περιεχόμενό της* (μέσα στο δικό του σύστημα πεποιθήσεων), *ψευδής όμως ως προς το ευρύ περιεχόμενό της* (σε σχέση με τον πραγματικό κόσμο έξω από αυτόν). Όταν όμως είχε την πεποίθηση ότι είναι δίκαιο να υπακούει στο νόμο του δικαστηρίου, αυτή ήταν *αληθής ως προς το στενό περιεχόμενό της*, ως προς το ευρύ όμως *δεν έχει νόημα να τη χαρακτηρίσουμε αληθή ή ψευδή*. Δεν έχουμε επιστημονικό-λογικό κριτήριο αλήθειας ή ψεύδους. Το κριτήριό μας είναι μια δική μας πεποίθηση, αν ο Σωκράτης είχε δίκιο ή άδικο, που και αυτή δεν επικυρώνεται λογικά. Ακόμα κι αν επικυρώνεται με λογικό συμπερασμό από μια ηθική αρχή (π.χ. την καντιανή προτροπή «πράττε έτσι ώστε οι πράξεις σου να μπορούν να αποτελέσουν καθολικό ηθικό νόμο»), η ηθική αυτή αρχή αφετηρίας δεν επικυρώνεται λογικά. Η προσπάθεια αναγωγής αυτών των ηθικών ή άλλων αξιών σε φυσικούς νόμους ή αρχές (π.χ. η απόδοση σε αυτές προσαρμοστικής αξίας στο πλαίσιο της εξελικτικής θεωρίας) συνιστά την αποκαλούμενη *νατουραλιστική πλάνη* (naturalistic fallacy) (βλ. Καφετζόπουλος 2005).

Μπορεί λοιπόν στο μέλλον να βρούμε πού εντοπίζονται οι αναπαραστάσεις αυτών των πεποιθήσεων, να προσδιορίσουμε το ελάχιστο νευρωνικό τους αντίστοιχο, να ανιχνεύσουμε τις μνημονικές εγγραφές που σχετίζονται με αυτές, καμία όμως νευροεπιστημονική τεχνική ή θεωρία δεν μπορεί να αποκαλύψει το σημασιολογικό τους περιεχόμενο. Μπορεί, ωστόσο, να μας αποκαλύψει αρκετά για τη λειτουργία τους, εφόσον οι αναπαραστάσεις αυτές υπάρχουν μέσα στον εγκέφαλο και έχουν μια σημαντική λειτουργία για τον οργανισμό. Η λειτουργία αυτή δεν μπορεί να είναι άλλη από την παραγωγή μιας όσο το δυνατόν καλύτερης ερμηνείας του περιβάλλοντος –υπό το φως προηγούμενων εμπειριών– και τη χρήση της στη λήψη αποφάσεων, στον προγραμματισμό της συμπεριφοράς και στη γλώσσα. Η λειτουργία αυτή, επίσης, δεν μπορεί παρά να ταυτίζεται με τη λειτουργία των νευρώνων του νευρωνικού υποστρώματος των αντίστοιχων αναπαραστάσεων (βλ. Crick and Koch 1998). Το φαινομενικό (στενό-υποκειμενικό) περιεχόμενο των αναπαραστάσεων που περιέχουν

¹³ Γι' αυτό πολλοί, αντί του όρου αναπαραστάση, προτιμούν τον όρο «εγγράμματα» (engrams) ή ακόμη και τον όρο «εξωγράμματα» (exograms) του Merlin Donald (1991, 2001).

¹⁴ Η Ε. Σαββάκη στην εισήγησή της στον πρώτο κύκλο υποστήριξε ότι η προσβαίνουσα συνείδηση αποτελεί το «εύκολο πρόβλημα» στην επιστημονική μελέτη της συνείδησης, επειδή, πράγματι, αφορά υπολογιστικού τύπου διαδικασίες που μπορούν να υλοποιηθούν ακόμη και σε έναν ηλεκτρονικό υπολογιστή, ενώ το «δύσκολο πρόβλημα» αφορά τη φαινομενική εμπειρία και την υποκειμενική διάσταση. Εδώ, ωστόσο, επιμένουμε ότι το «δύσκολο πρόβλημα» δεν είναι η υποκειμενική φύση της συνειδητής εμπειρίας αλλά το πρόβλημα του περιεχομένου των συνειδητών αναπαραστάσεων, που στην περίπτωση της προσβαίνουσας συνείδησης είναι σημαντικότερο, επειδή στον έλεγχο της συμπεριφοράς χρησιμοποιούνται αναπαραστάσεις που η σημασία τους δεν εξαντλείται σε όρους αλήθειας.

αισθητικές εμπειρίες (π.χ. «έχω την πεποίθηση ότι η ντομάτα –που βλέπω– είναι κόκκινη») παράγεται (επιγίνεται) από τη λειτουργία αυτών των νευρώνων και τις συνδέσεις τους.

Το περιεχόμενο όμως αναπαραστάσεων που περιέχουν πεποιθήσεις του τύπου «η Γη –που δεν βλέπω– είναι επίπεδη» δεν παράγεται μέσα στον εγκέφαλο, δεν επιγίνεται σε κάποια νευρωνική δομή, αλλά «εγγράφεται» απ' ευθείας μέσω της γλώσσας. Βρίσκεται σε έναν χώρο που θα διακινδυνεύαμε να τον ονομάσουμε «εκτεταμένη» ή «συλλογική» συνείδηση (κατ' αναλογία του *Extended Mind* των Clark and Chalmers [1998]) και μεταδίδεται μέσω διαφόρων κωδίκων από ομιλητή σε ομιλητή, από δάσκαλο σε μαθητή, από ειδικό σε μη ειδικό, κλπ. Οι εσωτερικές αναπαραστάσεις αυτών των πεποιθήσεων έχουν μια αιτιακή-ιστορική σχέση με τις εκτεταμένες (ευρείες) σημασίες και οικοδομούνται πιθανόν με τον ίδιο τρόπο που οικοδομούνται οι φυσικοί όροι της γλώσσας (βλ. Putnam 1975, σελ. 201 κ.ε.). Αυτή η ιστορικότητα και η διαχρονικότητα του αναπαραστασιακού περιεχομένου έχει οδηγήσει την Millikan (1993, σελ. 93) να δηλώσει ότι: «δεν υπάρχει τρόπος να εξεταστεί μόνο κατά την παρούσα στιγμή ένας άνθρωπος, π.χ. οι τρέχουσες ομιλιακές του προδιαθέσεις ή οι τρέχοντες σχηματισμοί του νευρωνικού δικτύου του, ούτως ώστε να αποκαλυφθεί έστω η αποβλεπτική φύση των εκφερόμενων προτάσεων ή των εσωτερικών αναπαραστάσεών του, πολύ περισσότερο να αποκαλυφθεί τι αυτές αναπαριστούν» (η έμφαση δική μου).

Καταλήγουμε λοιπόν στη ρήση του Fodor ότι η (νευρο)επιστήμη μπορεί να μας απαντήσει στο πώς αναπαρίστανται και πώς χρησιμοποιούνται οι συνειδητές καταστάσεις στον εγκέφαλο, δεν μπορεί να μας πει όμως από πού προέρχεται η σημασία τους. Η χαρά του Σωκράτη, ως συγκινησιακή αντίδραση, μπορεί να αναχθεί στη λειτουργία των κυκλωμάτων ευχαρίστησης του εγκεφάλου του, ενώ το φαινομενικό (στενό) της περιεχόμενο (το ότι αισθάνεται σαν να χαίρεται) μπορεί, αργά ή γρήγορα, να αναχθεί στη λειτουργία του νευρωνικού υποστρώματος των αναπαραστάσεων που την προκαλούν. Ως φαινομενική εμπειρία και ως συμπεριφορά παράγεται όπως η λύρα παράγει μουσική –η μουσική ως ήχος επιγίνεται στις φυσικές ιδιότητες και τη λειτουργία της λύρας. Το ευρύ ωστόσο περιεχόμενο των αναπαραστάσεων που οδηγούν στη χαρά του Σωκράτη –τη στιγμή μάλιστα που πρόκειται να πεθάνει–, η σημασία της αναπαράστασης του θεού του, των ηθικών ή άλλων αξιακών του πεποιθήσεων, δεν βρίσκεται μέσα στον εγκέφαλό του και δεν «παράγεται» από αυτόν, όπως η σημασία της μουσικής (η αισθητική της αξία) δεν παράγεται από τη λύρα που την παίζει. Ήταν προϊόν της ζωής του, της ιστορίας του, των σχέσεών του με άλλους ανθρώπους και εξακολουθεί, μετά το θάνατό του, να υπάρχει δίνοντας περιεχόμενο σε κρίσεις, πεποιθήσεις και επιθυμίες που καθορίζουν και ελέγχουν («νοσηματοδοτούν») ακόμη και τη δική μας συμπεριφορά.

Οποιοδήποτε λοιπόν αναγωγιστικό πρόγραμμα της προέλευσης ή της λειτουργίας της συνείδησης (και της συμπεριφοράς) δεν μπορεί να περιλάβει το ευρύ περιεχόμενο συνειδητών καταστάσεων που αναφέρονται σε αξιολογικές κρίσεις. Ο φόβος της αναγωγής οφείλεται σε αυτόν ακριβώς το φόβο μήπως οι επιστημονικές ερμηνείες εξαλείψουν το ευρύ (το βαθύ) περιεχόμενο της χαράς ή της λύπης μας, μήπως κάνουν περιττή την αναφορά μας στη συνείδηση όταν αναλύουμε τη συμπεριφορά μας, μήπως μας πείσουν ότι δεν είμαστε τίποτα παραπάνω από τα *zombies*, αυτά τα πλάσματα που μοιάζουν με ανθρώπους αλλά δεν διαθέτουν συνειδητές εμπειρίες και που έχουν εφεύρει οι φιλόσοφοι για τα νοητικά τους πειράματα. Ο φόβος αυτός είναι όχι μόνον

ανυπόστατος αλλά πρέπει να παραμεριστεί εντελώς, ώστε τα ερωτήματα στα οποία μπορεί να απαντήσει η επιστήμη να αποφορτιστούν ιδεολογικά και να οδηγήσουν σε ευρύτερα αποδεκτές περιγραφές και ερμηνείες, καθώς και σε γενικές θεωρίες που να αποτυπώνουν με σαφήνεια το σημερινό επίπεδο των γνώσεών μας.

Βιβλιογραφία

- Block N. (1995) On a confusion about a function of consciousness. *Behavioral and Brain Sciences* **18**: 227-247.
- Block N. (2005) Two neural correlates of consciousness. *Trends in Cognitive Sciences* **9**: 46-52. www.nyu.edu/gsas/dept/philo/faculty/block/papers/final_revised_proof.pdf
- Block N. and Stalnaker R. (1999) Dualism, conceptual analysis, and the explanatory gap. *Philosophical Review* **108**: 1-46.
- Brentano F. (1874/1973) *Psychology from an Empirical Standpoint*, (transl. A. Pan-curello, D. Terrel and L. McAllister). New York: Humanity Press.
- Chalmers D.J. (1995) The puzzle of conscious experience. *Scientific American* **273**: 62-68. consc.net/papers/puzzle.html
- Chalmers D.J. (1998) On the search for the neural correlate of consciousness, in Hamer-off S., Kaszniak A. and Scott A. (eds) *Toward a Science of Consciousness II: The Second Tucson Discussions and Debates*. Cambridge: MIT Press. <http://consc.net/papers/ncc.html>
- Chalmers D.J. (2000) What is a neural correlate of consciousness? In Metzinger T. (ed.) *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. Boston, Mass: MIT Press. <http://consc.net/papers/ncc2.html>.
- Clark A. and Chalmers D.J. (1998) The extended mind. *Analysis* **58**: 10-23. <http://consc.net/papers/extended.html>
- Crick F. (1994) *An Astonishing Hypothesis*. New York: Scribner's (ελλ. έκδοση *Μια εκπληκτική υπόθεση*, Αθήνα: Κάτοπτρο).
- Crick F and Koch C. (1990) Towards a neurobiological theory of consciousness. *Seminars in Neuroscience* **2**: 263-275.
- Crick F and Koch C. (1995) Why neuroscience may be able to explain consciousness, *Scientific American* **273**: 66-67.
- Crick F. and Koch C. (1998) Consciousness and neuroscience. *Cerebral cortex*, **8**: 97-107. www.klab.caltech.edu/~koch/crick-koch-cc-97.html
- Davidson D. (1980) *Essays on Actions and Events*. Oxford: Clarendon Press.
- Donald M. (1991) *Origins of the Modern Mind*. Cambridge, Mass.: Harvard University Press.
- Donald M. (2001) *A Mind So Rare: The Evolution of Human Consciousness*. New York: W.W. Norton.
- Dretske F. (1981) *Knowledge and the Flow of Information*. Cambridge, Mass: MIT Press.
- Eckhorn R., Bauer R., Jordan W., Brosch M., Kruse W., Munk M., and Reitboeck H.J. (1988) Coherent oscillations: a mechanism of feature linking in the visual cortex? *Biology and Cybernetics* **60**: 121-130.
- Edelman G.M. (1989) *The Remembered Present: A Biological Theory of Consciousness*. New York: Basic Books.
- Feigl H. (1958) The "Mental" and the "Physical", in Feigl H., Scriven M., and Maxwell G. (eds) *Concepts, Theories, and the Mind-Brain Problem (Minnesota Studies in the Philosophy of Science 2)*. Minnesota: University of Minnesota Press.
- Fodor J. (1987) *Psychosemantics – The Problem of Meaning in the Philosophy of Mind*. Cambridge, Mass: A Bradford Book, MIT Press.

- Fodor J. (1990) *A Theory of Content and Other Essays*. Cambridge, Mass: A Bradford Book, MIT Press.
- Gray C.M. and Singer W. (1989) Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proceedings of the National Academy of Sciences of USA* **86**: 1698-1702.
- Heath, R.G. (1964) Pleasure response of human subjects to direct stimulation of the brain: Physiologic and psychodynamic considerations, in R.G. Heath (ed.), *The Role of Pleasure in Human Behavior*, (pp. 219-243). New York: Hoeber.
- Heeger D.J., Boynton G.M., Demb J.B., Seidemann E., and Newsome W.T. (1999) Motion opponency in visual cortex. *Journal of Neuroscience* **19**: 7162-7174.
- Huk A.C., Ress D., and Heeger D.J. (2001) Neuronal basis of the motion aftereffect reconsidered. *Neuron* **32**: 161-172.
- Koch C. (2004) *The Quest of Consciousness*. Englewood: Roberts and Company Publishers.
- Koob G.F. and Le Moal M. (1997) Drug abuse: Hedonic omeostatic dysregulation. *Science* **278**: 52-58.
- Lange C.G. and James W. (1922) *The Emotions*, Baltimore: Williams and Wilkins.
- Leopold D.A. and Logothetis N.K. (1996) Activity changes in early visual cortex reflect monkeys' percept during binocular rivalry. *Nature* **379**: 549-553.
- Levine J. (1983) Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly* **64**: 354-361
- Libet B. (1993) The neural time factor in conscious and unconscious events, in *Experimental and Theoretical Studies of Consciousness (Ciba Foundation Symposium)*, New York: Willey.
- Llinás R. (2001), *I of the Vortex: From Neurons to Self*, Cambridge, Mass.: The MIT Press.
- Logothetis N. and Schall G. (1989) Neuronal correlates of subjective visual perception. *Science* **245**: 761-763.
- Lycan, W. (2005) Representational theories of consciousness, in Zalta E.N. (ed.), *The Stanford Encyclopedia of Philosophy (Spring 2005 On Line Edition)*: URL = <http://plato.stanford.edu/archives/spr2005/entries/consciousness-representational/>
- McGinn C. (1989) *Mental Content*. New York: Blackwell.
- Millikan G.R. (1984) *Language, Thought and Other Biological Categories – New Foundations for Realism*. Cambridge, Mass: A Bradford Book, MIT Press.
- Millikan G.R. (1993) *White Queen Psychology and Other Essays for Alice*. Cambridge, Mass: A Bradford Book, MIT Press.
- Nagel E. (1961) *The Structure of Science*. New York: Harcourt, Brace, and World.
- Olds, J. & Milner, P. (1954) Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology* **47**: 419-427.
- Place U.T. (1956) Is consciousness a brain process? *British Journal of Psychology* **47**: 44-50.
- Putman H. (1975) The meaning of “meaning”, in *Mind, Language and Reality. Philosophical Papers*, Vol. 2, pp. 215-271. Cambridge: Cambridge University Press.
- Rees G., Kreiman G. and Koch C. (2002) Neural correlates of consciousness in Humans. *Nature Reviews Neuroscience* **3**: 261-270.

- Roskies A.L. (1999) The binding problem. *Neuron* **24**: 7-9.
- Schachter S. and Singer J.E. (1962) Cognitive, social, and physiological determinants of emotional state. *Psychological Review* **69**: 379-399.
- Smart J.J.C. (1959) Sensations and brain processes. *Philosophical review* **68**: 141-156.
- Steinberg D.L. and Logothetis N.K. (1997) The role of temporal cortical areas in perceptual organization. *Proceedings of the National Academy of Sciences USA* **94**: 3408-3413.
- Sutton J. (2004) Representation, reduction, and interdisciplinarity in the sciences of memory, in Clapin H., Staines P., and Slezak P. (eds.), *Representation in Mind: New Approaches in Mental Representation*, pp. 187-216. New York: Greenwood Publishers.
- http://www.phil.mq.edu.au/staff/jsutton/Sutton_Representation_Memory.PDF