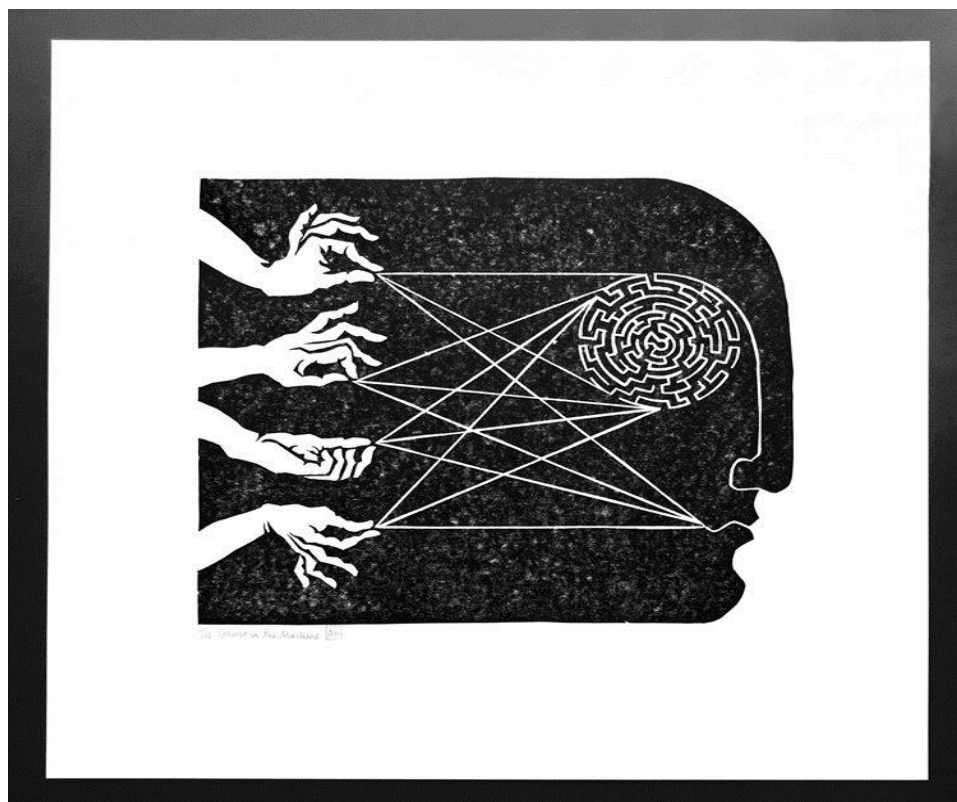




Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών
Τομέας Ανθρωπιστικών, Κοινωνικών Επιστημών και Δικαίου

Μελέτη για την υπολογισσιμότητα των ενσυνείδητων λειτουργιών: Διάλογος Roger Penrose – Paul Churchland



Διπλωματική Εργασία

Κόττας Χουλάκης Γιάννης

Επιβλέπων:

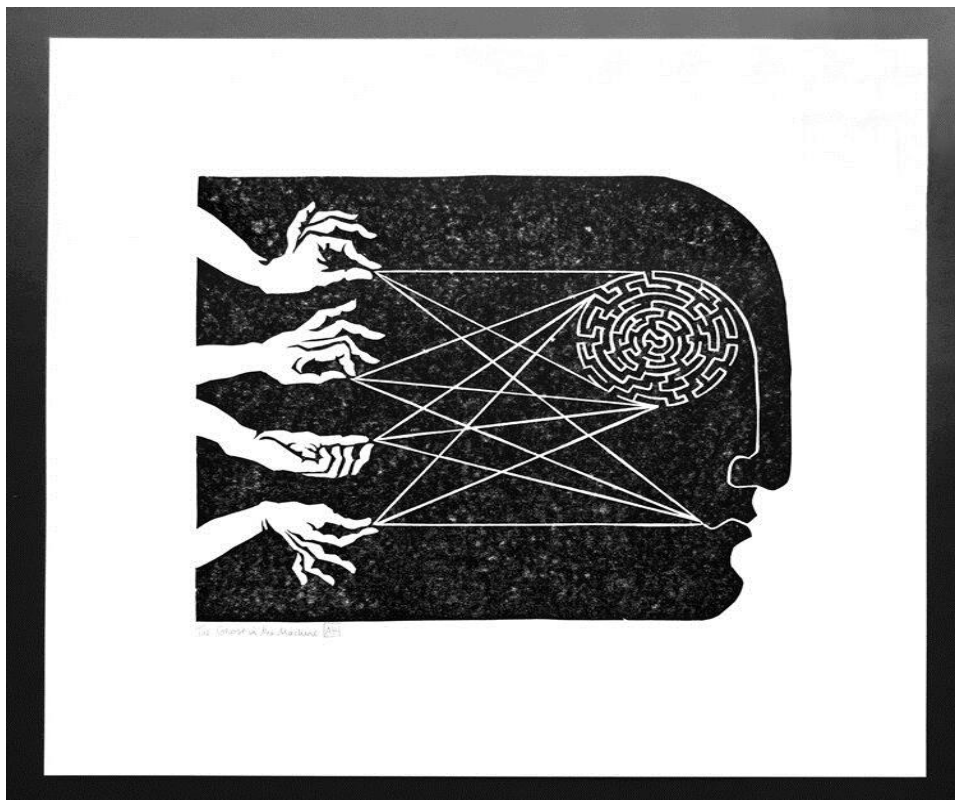
Στέλιος Σπυρίδων, μέλος ΕΔΙΠ

Αθήνα 2023



National Technical University of Athens
School of Applied Mathematical and Physical Science
Department of Humanities, Social Sciences and Law

A study on the computability of conscious functions: Roger Penrose - Paul Churchland Discourse



Diploma Thesis

Kottas Choulakis Ioannis

Supervisor:

Stelios Spyridon, Teaching and Research Associate

Athens 2023

Υπεύθυνη Δήλωση εναντίον λογοκλοπής

Έχω διαβάσει και κατανοήσει τους κανόνες για τη λογοκλοπή και τον τρόπο σωστής αναφοράς των πηγών που περιέχονται στον Οδηγό συγγραφής Διπλωματικών Εργασιών. Δηλώνω ότι, από όσα γνωρίζω, το περιεχόμενο της παρούσας εργασίας είναι προϊόν δικής μου δουλειάς και υπάρχουν αναφορές σε όλες τις πηγές που χρησιμοποίησα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτή τη Διπλωματική εργασία είναι του συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις της Σχολής Μηχανολόγων Μηχανικών ή του Εθνικού Μετσόβιου Πολυτεχνείου.

Ονοματεπώνυμο

Κόττας Χουλάκης Γιάννης

Ευχαριστίες

Με την παρούσα διπλωματική εργασία ολοκληρώνονται αισίως οι προπτυχιακές σπουδές μου στη Σχολή των Μηχανολόγων Μηχανικών, του Εθνικού Μετσόβιου Πολυτεχνείου.

Δράττομαι της ευκαιρίας για να ευχαριστήσω όλους τους ανθρώπους που στάθηκαν δίπλα μου, με όποιο τρόπο, όλα αυτά τα χρόνια, και με βοήθησαν να ανακαλύψω τον εαυτό μου και να κυνηγήσω τους στόχους μου. Αυτά τα χρόνια, με τις αναποδιές και τις χαρές τους, χαρακτήρισαν ανεξίτηλα τη μετεφηβική και την ενήλικη, ως τώρα, ζωή μου.

Από τις καθηγήτριες και τους καθηγητές που καταφέρνουν να σπάσουν το επικοινωνιακό φράγμα μεταξύ διδάσκοντα και διδασκόμενου, το διδακτικό προσωπικό που αφιερώνει το χρόνο του στην παραγωγή και μετάδοση επιστημονικής και χρήσιμης, για την κοινωνία, γνώσης, μέχρι τους συμφοιτητές που ξημεροβραδιαζόμασταν παρέα πάνω από σημειώσεις, και τις φίλες που μοιραστήκαμε μοναδικές στιγμές εντός και εκτός αμφιθεάτρων.

Ιδιαίτερα επιθυμώ να ευχαριστήσω τον διδάσκοντα της Φιλοσοφίας των Επιστημών στο ΕΜΠ και επιβλέποντα της παρούσας διπλωματικής εργασίας, Σπυρίδωνα Στέλιο και τον αναπληρωτή καθηγητή του Τομέα Μαθηματικών του ΕΜΠ, Πέτρο Στεφανέα, για την επιστημονική και συμβουλευτική καθοδήγηση που μου προσέφεραν σε όλα τα στάδια εκπόνησης της εργασίας με τις εύστοχες και εποικοδομητικές παρατηρήσεις τους.

Η ολοκλήρωση αυτής της εργασίας θα ήταν αδύνατη χωρίς την αμέριστη βοήθεια και συμπαράσταση του ομότιμου καθηγητή ΕΜΠ και ανήσυχου πνεύματος, Γεράσιμου Σπαθή. Τόσο για τη συμβολή του στη σύλληψη της ιδέας της εργασίας, για τις διεξοδικές μας συζητήσεις σχετικά με το αντικείμενο της μελέτης, όσο και για τον τρόπο με τον οποίο στάθηκε δίπλα μου, πνευματικά και ηθικά, όλον αυτόν τον καιρό.

Τέλος, θα ήθελα να ευχαριστήσω βαθιά το Θεοδωρή, τη Σοφία και το Χάρη, για την άνευ όρων αγάπη και στήριξη που μου προσφέρουν, εξ απαλών ονύχων.

Κόττας Χουλιάκης Γιάννης, 10/10/2023

Περιεχόμενα

Ευχαριστίες	2
Περιεχόμενα	3
Περίληψη	5
Abstract	6
Κεφάλαιο 1 : Εισαγωγή	7
1.1 Φιλοσοφικό πλαίσιο.....	7
1.2 Βασικό πρόβλημα	10
1.3 Στόχοι της εργασίας.....	11
1.4 Μεθοδολογία.....	12
Κεφάλαιο 2 : Το πρόβλημα νου-σώματος	14
2.1 Ο δυϊσμός του Descartes.....	14
2.2 Ο μονισμός του Spinoza	16
2.4 Θεωρία Ταυτότητας Νου-Εγκεφάλου.....	18
2.5 Λειτουργισμός.....	19
2.6 Η κριτική του Gilbert Ryle στο καρτεσιανό δόγμα.....	21
Κεφάλαιο 3 : Η φιλοσοφική ματιά του Roger Penrose στα ερωτήματα της συνείδησης	26
3.1 Εισαγωγή στη σκέψη του R.Penrose	26
3.2 Μη υπολογιστική προσέγγιση της συνείδησης	27
3.2.1 Σχετικά με το δυϊσμό σώματος-νου	28
3.3 Το θεώρημα του Gödel	29
3.4 Η χρονική ροή στον Penrose	30
3.5 Τα όρια της Τεχνητής Νοημοσύνης	32
3.6 Η νέα φυσική που χρειαζόμαστε για την εξήγηση της νόησης	35
3.7 Κριτική και αξιολόγηση των θέσεων του Penrose	37

3.7.1 Η κριτική του Paul Churchland.....	37
Κεφάλαιο 4 : Η φιλοσοφική ματιά του Paul Churchland στα ερωτήματα της συνείδησης.....	40
4.1 Εισαγωγή στη σκέψη του P.Churchland.....	40
4.2 Η κριτική του Churchland στο Turing τεστ.....	40
4.3 Η υπολογιστική θεωρία της συνείδησης.....	42
4.3.1 Ο ρόλος των νευρωνικών δικτύων στην ανθρώπινη νοημοσύνη	43
4.3.2 Τα θεμελιώδη χαρακτηριστικά της συνείδησης	45
4.4 Η γλωσσική λειτουργία στον Churchland	51
4.4.1 Κριτική στην έννοια του γλωσσικού οργάνου.....	51
4.4.2 Κριτική στη γλωσσοκεντρική θεωρία του Dennett	52
4.4.3 Η σημασία της γλώσσας στη σκέψη και στη συνείδηση	54
4.5 Η εννοιολογική διεύρυνση και η αλλαγή παραδείγματος	55
4.5.1 Η κατανόηση της θεωρίας μέσω «παραδειγμάτων».....	55
4.5.2 Ο ρόλος του «παραδείγματος» στη νευροϋπολογιστική θεωρία του Churchland.....	58
4.5.3 Η εννοιολογική επέκταση στο έργο του Αριστείδη Μπαλτά.....	59
4.6 Το πείραμα του Κινέζικου Δωματίου και τα qualia	60
4.6.1 Ο Searle μέσα στο δωμάτιο.....	61
4.6.2 Οι αισθητηριακές ποιότητες εντός του λειτουργισμού.....	63
4.6.3 Πώς είναι να είσαι;.....	64
Κεφάλαιο 5 : Εναλλακτικές προσεγγίσεις	67
5.1 Η θεωρία του Daniel Dennett	67
5.2 Η Ολοκληρωμένη Θεωρία Πληροφοριών	70
Κεφάλαιο 6 : Με το βλέμμα στο μέλλον.....	73
6.1 Σύνοψη των ευρημάτων.....	73
6.1.1 Κοινοί τόποι Penrose – Churchland.....	74
6.1.2 Διαχωριστικές γραμμές Penrose – Churchland.....	75
6.2 Μελλοντικές κατευθύνσεις	76
Κατάλογος Εικόνων	78
Βιβλιογραφία	79

Περίληψη

Η εργασία επιχειρεί να παρουσιάσει και να σχολιάσει ορισμένα κεντρικά ζητήματα σε σχέση με τη φύση της συνείδησης, με τη μορφή συνομιλίας μεταξύ του Roger Penrose και του Paul Churchland. Ο Penrose, διάσημος μαθηματικός και φυσικός, στο θεμελιώδες έργο του «Σκιές του Νου» (Shadows of the Mind) υποστηρίζει μια μη υπολογιστική θεωρία της συνείδησης, προτείνοντας μια κβαντομηχανική βάση για την ανθρώπινη νόηση. Αντίθετα, ο Paul Churchland, διακεκριμένος φιλόσοφος του νου, παρουσιάζει μια υπολογιστική θεωρία της συνείδησης στο βιβλίο του «Η μηχανή της Λογικής, η Θέση της Ψυχής» (The Engine of Reason, the Seat of the Soul), δίνοντας έμφαση στο ρόλο των νευρωνικών δικτύων και της μαζικά παράλληλης αρχιτεκτονικής του εγκεφάλου. Σκοπό της εργασίας αποτελεί η κριτική εξέταση των προοπτικών και των επιχειρημάτων που προβάλλουν ο Penrose και ο Churchland και η αξιολόγηση των επιπτώσεών τους στην κατανόηση της συνείδησης. Μέσω της ενδελεχούς εξέτασης των βασικών εννοιών και των λογικών πλαισίων που αναπτύσσονται στα κείμενά τους, παρουσιάζονται οι κοινοί τόποι, αλλά και οι διαχωριστικές γραμμές μεταξύ των δύο συγγραφέων. Σημαντικό κομμάτι της εργασίας καταλαμβάνει η κριτική που έχει ασκηθεί απέναντι στους δυο διανοητές από την υπόλοιπη επιστημονική κοινότητα, αλλά και η συνοπτική παρουσίαση διαφορετικών συγγενικών θέσεων. Αξιολογείται, ακόμα, η επιστημονική ορθότητα και η εμπειρική τεκμηρίωση των δυο θεωρήσεων σε σχέση με τις σύγχρονες επιστημονικές εξελίξεις. Επιπλέον, δίνεται βάση στην ανάλυση των συνεπειών των θεωριών τους στις επικρατούσες φιλοσοφικές τάσεις τόσο σε σχέση με το δυϊσμό νου-σώματος όσο και σε σχέση με τη δυνατότητα μηχανών τεχνητής νοημοσύνης για ενσυνείδητη αντίληψη. Μέσα από την παρουσίαση και αντιπαράθεση των δυο θέσεων, επιχειρείται να αναδειχθεί η σχέση των έργων τους με την προοπτική της ριζικής εμμένειας.

Abstract

This paper attempts to present and comment on some central issues in relation to the nature of consciousness, in the form of a conversation between Sir Roger Penrose and Paul Churchland. Penrose, a renowned mathematician and physicist, in his seminal work *Shadows of the Mind* argues for a non-computational theory of consciousness, proposing a quantum mechanical basis for human cognition. In contrast, Paul Churchland, a prominent philosopher of mind, presents a computational theory of consciousness in his book *The Engine of Reason, the Seat of the Soul*, emphasizing the role of neural networks and the massively parallel architecture of the brain. The aim of the paper is to critically examine the perspectives and arguments put forward by Penrose and Churchland and to assess their implications for our understanding of consciousness. Through a thorough examination of the key concepts and logical frameworks developed in their writings, common ground as well as dividing lines between the two authors are presented. An important part of the paper is taken up by the criticisms that have been made of the two thinkers by the rest of the scientific community, as well as by the brief presentation of different related positions. In addition, attention is given to the analysis of the implications of their theories for prevailing philosophical trends both in relation to the mind-body dualism and in relation to the potential of AI machines for conscious perception. Through the presentation and confrontation of the two positions, an attempt is made to highlight the relationship between their works and the perspective of radical immanence.

Κεφάλαιο 1 : Εισαγωγή

1.1 Φιλοσοφικό πλαίσιο

Η συνείδηση, η υποκειμενική εμπειρία της επίγνωσης του εαυτού μας και του κόσμου γύρω μας, παραμένει ένα από τα πιο γοητευτικά και διαρκή μυστήρια της ανθρώπινης ύπαρξης. Καθώς εμβαθύνουμε στη μελέτη της συνείδησης, συναντάμε αμέτρητα βαθιά και θεμελιώδη ερωτήματα που προκαλούν την κατανόησή μας για τη φύση και την προέλευση αυτού του αινιγματικού φαινομένου.

Ακολουθούν μερικά επίκαιρα ερωτήματα με επίκεντρο τη συνείδηση:

- Μπορεί να αποτελεί αντικείμενο επιστημονικής μελέτης η πνευματική μας υπόσταση με τις αντίστοιχες νοητικές της διαδικασίες ή πρέπει οι επιστημονικές έρευνες να περιοριστούν στην μελέτη των υλικών χαρακτηριστικών του σύμπαντος; Μοιάζει εφικτή μια μελλοντική επιστημονική εξήγηση του μυστηρίου της νόησης;
- Ποια είναι η σχέση μεταξύ της συνείδησης και της επεξεργασίας πληροφοριών στην τεχνητή νοημοσύνη; Καθώς οι τεχνολογίες τεχνητής νοημοσύνης και μηχανικής μάθησης εξελίσσονται, πώς θα καθορίσουμε αν και πότε οι μηχανές διαθέτουν συνείδηση ή εμπειρίες που μοιάζουν με συνείδηση;
- Μπορεί να υπάρχει συνείδηση σε μη βιολογικές οντότητες;
- Πώς αναδύεται η συνείδηση κατά την ανάπτυξη του ανθρώπινου εγκεφάλου; Ποιοι είναι οι συγκεκριμένοι μηχανισμοί και διεργασίες που γεννούν την υποκειμενική εμπειρία καθώς ο εγκέφαλος αναπτύσσεται από τη βρεφική ηλικία έως την ενηλικίωση;
- Είναι η συνείδηση ένα συνεχές ή διακριτό φαινόμενο; Υπάρχει η συνείδηση ως συνεχής ροή εμπειρίας ή αποτελείται από διακριτές στιγμές ή "στιγμιότυπα" που δίνουν την ψευδαίσθηση της συνέχειας;
- Ποια είναι η σχέση μεταξύ της συνείδησης και της ελεύθερης βούλησης; Μπορούν οι συνειδητές εμπειρίες μας να επηρεάσουν τις διαδικασίες λήψης αποφάσεων ή είναι απλά υποπροϊόντα ντετερμινιστικών νευρωνικών διεργασιών;
- Πώς ο πολιτισμός και η γλώσσα διαμορφώνουν τη συνείδηση; Σε ποιο βαθμό τα πολιτισμικά και γλωσσικά πλαίσια επηρεάζουν τις υποκειμενικές μας εμπειρίες και διαμορφώνουν την κατανόηση των συνειδητών λειτουργιών;
- Ποιο ρόλο παίζει η προσοχή στη διαμόρφωση της συνείδησης; Πώς η προσοχή επιλέγει και φιλτράρει τις πληροφορίες, επηρεάζοντας τι εισέρχεται στη συνειδητή μας επίγνωση και τι παραμένει εκτός αυτής;
- Μπορούμε να δημιουργήσουμε μια ενιαία θεωρία της συνείδησης που να περιλαμβάνει τόσο τις υποκειμενικές πτυχές της εμπειρίας όσο και τις υποκείμενες νευρωνικές και υπολογιστικές διαδικασίες;

Τα ερωτήματα αυτά αντικατοπτρίζουν τις συνεχιζόμενες εξερευνήσεις και συζητήσεις στον τομέα των μελετών της συνείδησης, αναδεικνύοντας τη δυναμική φύση

της έρευνας και τα βαθιά μυστήρια που συνεχίζουν να πυροδοτούν αντιπαραθέσεις μεταξύ ερευνητών και φιλοσόφων.

Προφανώς, τα ερωτήματα έχουν περισσότερο ρητορικό χαρακτήρα και σκοπό έχουν να υπογραμμίσουν το μέγεθος και τη δυσκολία του προβλήματος. Η αναζήτηση ολοκληρωμένων και καταληκτικών απαντήσεων στους παραπάνω επίκαιρους προβληματισμούς ξεπερνά κατά πολύ τα καθήκοντα της παρούσας εργασίας. Θα αποπειραθούμε, ωστόσο, να αναμετρηθούμε με συγκεκριμένες πτυχές των προβληματικών που παρουσιάστηκαν.

Στην παρούσα εργασία, εξερευνούνται διαφορετικές οπτικές για το φαινόμενο της συνείδησης και τη δυνατότητα υπολογισμού των λειτουργιών της, σε μια προσπάθεια συγκερασμού των παραδόσεων της φιλοσοφίας του νου και των εμπειρικών μεθόδων της επιστημονικής σκέψης. Κεντρικό άξονα της εργασίας αποτελεί η υποθετική ‘συνομιλία’ και ανταλλαγή επιχειρημάτων από δύο επιφανείς στοχαστές, τον Roger Penrose και τον Paul Churchland, γύρω από τον οποίο αναπτύσσονται και σχολιάζονται διαφορετικές συλλογιστικές ερευνητών και επιστημόνων που έχουν επεξεργαστεί εκτενώς, διαχρονικά, αλλά και επίκαιρα, ερωτήματα που αφορούν τη φύση της συνείδησης.

Η φιλοσοφία του νου είναι ένας κλάδος της φιλοσοφίας που μελετά τη φύση του νου, τα νοητικά γεγονότα, τις νοητικές λειτουργίες και ιδιότητες, τη συνείδηση, και τη σχέση τους με το φυσικό σώμα, ιδιαίτερα τον εγκέφαλο. Το πρόβλημα νου-σώματος, δηλαδή η σχέση του νου με το σώμα, συνήθως θεωρείται ως ένα κεντρικό ζήτημα στη φιλοσοφία του νου.

Ακολουθεί μια συνοπτική αναφορά σε δυο κυρίαρχα φιλοσοφικά ρεύματα, ώστε να τεθεί το βασικό πλαίσιο γύρω από το οποίο θα εξελιχθεί η συζήτηση. Τα δυο αυτά ρεύματα αναπτύχθηκαν ως διαφορετικές προσεγγίσεις στο βασικό ερώτημα της φιλοσοφίας, τη σχέση δηλαδή, Είναι και νόησης, ύλης και συνείδησης. Από τη μία, οι εκπρόσωποι του *ιδεαλισμού* θεωρούν ότι η συνείδηση (το πνεύμα, η νόηση) προϋπήρξε χρονικά της ύλης. Ο υλικός κόσμος παράγεται και εξαρτάται από το άυλο πνεύμα. Από την άλλη, οι *υλιστές* φιλόσοφοι δίνουν προβάδισμα στην ύλη έναντι του πνεύματος. Ο άφθαρτος και αιώνιος υλικός κόσμος είναι ανεξάρτητος της συνείδησης, καθώς τα υλικά αντικείμενα έχουν την ιδιότητα να υπάρχουν έξω από την ανθρώπινη συνείδηση και να αντανακλώνται από αυτή. Πιο συγκεκριμένα, παρακάτω, παρουσιάζονται οι πτυχές των δυο ρευμάτων¹ που σχετίζονται με το ζήτημα της υπολογισιμότητας της συνείδησης.

- **Υλισμός-Φυσικαλισμός**

Οι σύγχρονες υλιστικές προσεγγίσεις υποστηρίζουν ότι η συνείδηση μπορεί να εξηγηθεί εξ ολοκλήρου από φυσικές διαδικασίες και οντότητες, όπως ο εγκέφαλος και οι νευρικές του δραστηριότητες. Η κατανόηση των φυσικών και βιολογικών μηχανισμών που

¹ Πρόκειται για μια πρόχειρη κατηγοριοποίηση, με σκοπό να τεθούν ορισμένες βασικές διαχωριστικές γραμμές ανάμεσα στις δυο σχολές σκέψης. Προφανώς, υλισμός και ιδεαλισμός αποτελούν γενικές έννοιες (umbrella terms) ή οικογένειες θεωριών οι οποίες περιλαμβάνουν μια σειρά από διαφορετικά παρακλάδια θέσεων και αναλύσεων. Στα επόμενα κεφάλαια αναπτύσσονται περαιτέρω συγκεκριμένες πτυχές τους.

διέπουν τη συνείδηση μπορεί να οδηγήσει τελικά σε μια ολοκληρωμένη εξήγηση αυτού του φαινομένου. Οι έρευνες συχνά βασίζονται σε εμπειρικά στοιχεία από τις νευροεπιστήμες, τη γνωστική ψυχολογία και άλλους συναφείς τομείς.

Ερευνητές που αποδέχονται κατ' αρχήν τον υλισμό, υιοθετούν συχνά την άποψη ότι η συνείδηση είναι μια αναδυόμενη ιδιότητα που προκύπτει από τις πολύπλοκες αλληλεπιδράσεις των νευρώνων και των νευρωνικών δικτύων στον εγκέφαλο. Υποστηρίζουν ότι η μελέτη της δομής και των λειτουργιών του εγκεφάλου, η εξέταση των νευρωνικών συσχετίσεων των ενσυνείδητων καταστάσεων και η διερεύνηση της επεξεργασίας πληροφοριών εντός του εγκεφάλου, μπορούν να φωτίσουν σημαντικές πλευρές της συνείδησης που αγνοούμε ως τώρα. Οι υποστηρικτές του υλισμού δίνουν προτεραιότητα στην χρήση επιστημονικών μεθοδολογιών και εμπειρικών παρατηρήσεων για την αποκάλυψη των υποκείμενων μηχανισμών της συνείδησης.

- **Ιδεαλισμός**

Σε αντίθεση με τον υλισμό, ο ιδεαλισμός, γενικά, θεωρεί ότι η συνείδηση δεν μπορεί να εξηγηθεί πλήρως μόνο με φυσικές διαδικασίες. Φιλόσοφοι και ερευνητές που αποδέχονται τον καρτεσιανό οντολογικό δυισμό², υποστηρίζουν την ύπαρξη μιας διακριτής νοητικής ή συνειδησιακής σφαίρας που είναι ξεχωριστή από τον φυσικό κόσμο. Ως συνέπεια, η συνείδηση δεν μπορεί να αναχθεί στις λειτουργίες του εγκεφάλου, αλλά αντίθετα συνίσταται σε μια άυλη ουσία ή υπόσταση που αλληλεπιδρά με το φυσικό σώμα. Οι ιδεαλιστές συχνά επικαλούνται φιλοσοφικά επιχειρήματα και εσωτερικές εμπειρίες για να υποστηρίξουν τους ισχυρισμούς τους. Μια βασική αντίρρηση στις υλιστικές αναγνώσεις αφορά τις υποκειμενικές συνειδητές εμπειρίες, όπως τα *qualia* (οι ακατέργαστες, υποκειμενικές ιδιότητες των αισθητηριακών εμπειριών), τα οποία δεν μπορούν να εξηγηθούν αποκλειστικά από φυσικές διεργασίες. Για τη ιδεαλιστική παράδοση η συνείδηση διαθέτει μοναδικές ιδιότητες που υπερβαίνουν τον υλικό κόσμο, καθιστώντας αναγκαία την ύπαρξη ενός πρόσθετου επεξηγηματικού πλαισίου πέρα από το πεδίο εφαρμογής του υλισμού.

Μια ακόμα κατηγοριοποίηση σε σχέση με τη φύση της πραγματικότητας και της ύπαρξης είναι αυτή μεταξύ μονισμού-δυισμού. Για το *δυισμό* η πραγματικότητα συνίσταται σε διακριτές ουσίες ή υποστάσεις. Η πιο διαδεδομένη μορφή δυισμού είναι ο *δυϊσμός ουσίας* (*substance dualism*), που υποστηρίζει την ύπαρξη δύο διακριτών ουσιών, του νου και του σώματος. Άλλες δυϊστικές προσεγγίσεις κυμαίνονται από τον *αλληλεπιδραστικό δυισμό* (*interactive dualism*), ο οποίος προϋποθέτει αμφίδρομη αλληλεπίδραση μεταξύ του φυσικού και του νοητικού κόσμου, έως και πιο διαφοροποιημένες παραλλαγές, όπως ο *δυϊσμός των ιδιοτήτων* (*property dualism*). Για τους *δυϊστές* αυτού του είδους η συνείδηση προκύπτει από συγκεκριμένες ιδιότητες ή πτυχές των φυσικών οντοτήτων, αλλά δεν αποτελεί ξεχωριστή ουσία. Ο *μονισμός*, αντίθετα με τον *δυισμό*, υποστηρίζει ότι υπάρχει μόνο μία βασική ουσία ή αρχή στον κόσμο. Ο *μονισμός διακριτών μορφών* περιλαμβάνει τον *φυσικό μονισμό* (*physical monism*), που υποστηρίζει την αναγωγή της ύπαρξης στις

²Στο επόμενο κεφάλαιο γίνεται εκτενέστερη αναφορά στο φιλοσοφικό έργο του Γάλλου στοχαστή Renes Descartes (1596-1650).

φυσικές διαδικασίες, και τον ιδεαλιστικό μονισμό (idealistic monism), που υποστηρίζει ότι η πραγματικότητα είναι ουσιαστικά νοητική.

Διαπιστώνουμε, λοιπόν, ότι η μελέτη της συνείδησης μάς επιφυλάσσει ένα ευρύ φάσμα σύνθετων και θεμελιωδών ερωτημάτων. Ενώ ο υλισμός δίνει έμφαση στις φυσικές και νευρωνικές δομές της συνείδησης, ο ιδεαλισμός αντιτείνει την ύπαρξη ενός άυλου συνειδησιακού πεδίου, διαχωρισμένου και ανεξάρτητου από την υλική πραγματικότητα. Και οι δύο προοπτικές προσφέρουν χρήσιμες προσεγγίσεις για την κατανόηση και την εξήγηση της συνείδησης, επικαλούμενες επιστημονικά στοιχεία, εμπειρικές παρατηρήσεις, φιλοσοφικούς συλλογισμούς και εσωτερικές εμπειρίες. Καθώς εμβαθύνουμε στην εξερεύνηση της συνείδησης, είναι σημαντικό να εμπλακούμε σε διεπιστημονικό διάλογο, να προωθήσουμε τη συνεργασία μεταξύ διαφορετικών απόψεων και να παραμείνουμε ανοιχτοί στις δυνατότητες που μπορεί να προκύψουν από τη μελλοντική έρευνα και τις επιστημονικές ανακαλύψεις. Η διαρκής ενασχόλησή μας -σε πρακτικό και θεωρητικό επίπεδο- με αυτά τα θεμελιώδη ερωτήματα, μας φέρνει όλο και πιο κοντά στην απομάγευση των μυστηρίων που σκιάζουν την κατανόησή μας για τη συνείδηση και τη φύση της υποκειμενικής εμπειρίας.

Η συζήτηση μεταξύ Penrose και Churchland λαμβάνει χώρα μέσα σε αυτό το ευρύτερο πλαίσιο. Ο Sir Roger Penrose, ένας διακεκριμένος μαθηματικός και φυσικός, προσφέρει τη δικιά του προοπτική στη συζήτηση εφαρμόζοντας γνώσεις από την κβαντομηχανική στο πρόβλημα της συνείδησης. Από την άλλη πλευρά, ο Paul Churchland, διακεκριμένος φιλόσοφος του νου και γνωστικός επιστήμονας, εστιάζει στις υπολογιστικές πτυχές της συνείδησης και στον ρόλο των νευρωνικών δικτύων. Και οι δύο προσπαθούν με τον τρόπο τους και από διαφορετική σκοπιά ο καθένας, να συμβάλλουν στη συνεχιζόμενη αναζήτηση μιας ολοκληρωμένης θεωρίας της συνείδησης. Τα επιχειρήματά τους αντανακλούν διαφορετικές φιλοσοφικές και επιστημονικές παραδόσεις, με τον Penrose να αντλεί από την εμπειρία του στη φυσική και τα μαθηματικά, ενώ ο Churchland δίνει έμφαση στη μελέτη των νευρωνικών μηχανισμών και των υπολογιστικών διαδικασιών.

Η κατανόηση του ιστορικού υποβάθρου και του πνευματικού πλαισίου γύρω από τη φύση της συνείδησης είναι ζωτικής σημασίας για την εκτίμηση της συμβολής των Penrose και Churchland. Στη συνέχεια της εργασίας θα εξεταστεί σε μεγαλύτερο βάθος το θεμελιώδες ερώτημα της φιλοσοφίας που αφορά τη σχέση νου-σώματος, μέσα από το πρίσμα ορισμένων διακεκριμένων φιλοσόφων και κεντρικών ερμηνειών.

1.2 Βασικό πρόβλημα

Το κεντρικό πρόβλημα με το οποίο έρχονται αντιμέτωποι οι Penrose και Churchland στα αντίστοιχα έργα τους, είναι η φύση της συνείδησης και η δυνατότητα ή μη υπολογισιμότητάς της μέσω αλγοριθμικών διεργασιών. Η συνείδηση είναι ένα σύνθετο φαινόμενο που περιλαμβάνει τις υποκειμενικές μας εμπειρίες, την αυτοεπίγνωση και τις αυτοαναφορικές σκέψεις μας. Η κατανόηση του τρόπου με τον οποίο η συνείδηση προκύπτει από τις φυσικές διεργασίες του εγκεφάλου αποτελεί σημαντική πρόκληση για την επιστημονική και φιλοσοφική κοινότητα.

Οι δυο επιστήμονες προσεγγίζουν το πρόβλημα αυτό από διαφορετικές φιλοσοφικές και επιστημονικές οπτικές γωνίες και οδηγούνται σε διαφορετικά συμπεράσματα. Ο Penrose τάσσεται κατά του λειτουργισμού και των υπολογιστικών μοντέλων της συνείδησης και αναζητά μια εναλλακτική εξήγηση που έχει τις ρίζες της στην κβαντομηχανική. Αμφισβητεί κατά πόσον η συνείδηση μπορεί να αναχθεί μόνο σε υπολογιστικούς αλγορίθμους ή νευρωνικά δίκτυα. Στόχος του Penrose είναι να παράσχει μια μη υπολογιστική εξήγηση της συνείδησης που να εξηγεί τον πλούτο της υποκειμενικής εμπειρίας.

Αντίθετα, ο Churchland προτείνει μια υπολογιστική θεωρία της συνείδησης, υποστηρίζοντας ότι ο εγκέφαλος λειτουργεί ως ένα εξελιγμένο σύστημα επεξεργασίας πληροφοριών. Δίνει έμφαση στο ρόλο των νευρωνικών δικτύων και της διανυσματικής κωδικοποίησης, της ενσωμάτωσης των αισθητηριακών εισροών στη δημιουργία συνειδητών εμπειριών. Στόχος του Churchland είναι να διαμορφώσει ένα πλαίσιο που να εξηγεί τη συνείδηση με βάση υπολογιστικές αρχές.

Η διαφωνία μεταξύ των Penrose και Churchland περιστρέφεται γύρω από τις θεμελιώδεις υποθέσεις τους σχετικά με τη φύση της συνείδησης και τα κατάλληλα επεξηγηματικά πλαίσια για την κατανόησή της. Με τη διερεύνηση των επιχειρημάτων τους και την εξέταση των στοιχείων και της συλλογιστικής πίσω από τις αντίστοιχες θέσεις τους, μπορούμε να αποκτήσουμε βαθύτερες γνώσεις για τη φύση της συνείδησης και τις πιθανές οδούς για τη διερεύνησή της.

1.3 Στόχοι της εργασίας

Οι κυριότεροι ερευνητικοί στόχοι της παρούσας εργασίας είναι να παρουσιάσει και να αναλύσει κριτικά τα επιχειρήματα των Roger Penrose και Paul Churchland όπως αναπτύσσονται στα αντίστοιχα έργα τους, «Σκιές του Νου» (Shadows of the Mind) και «Η Μηχανή της Λογικής, η Θέση της Ψυχής» (The Engine of Reason, the Seat of the Soul), σχετικά με την εξήγηση των φαινομένων των συνειδητών λειτουργιών.

Το παζλ των γνωστικών αντικειμένων με τα οποία συνδιαλέγεται η εργασία, συντίθεται από προσεγγίσεις που εδράζονται στα πεδία της Φιλοσοφίας του Νου, της Λογικής, της Επιστήμης των Υπολογιστών και της Τεχνητής Νοημοσύνης, των Γνωσιακών Επιστημών και των Νευροεπιστημών. Με κριτική διάθεση, επιχειρείται μια αδρή γενεαλογία ορισμένων φιλοσοφικών ρευμάτων σκέψης που στοχεύει στη διασάφηση εννοιών και την επεξήγηση του θεωρητικού πλαισίου, εντός του οποίου αρθρώνονται κάποια βασικά ζητήματα που απασχολούν τη σύγχρονη επιστημονική σκέψη, γύρω από τη φύση της συνείδησης.

Εξετάζοντας και αξιολογώντας τις θέσεις των στοχαστών, η παρούσα έρευνα αποσκοπεί στο:

- Να εντοπίσει τα κύρια επιχειρήματα και τις θεωρίες που διατυπώνουν οι Penrose και Churchland, αλλά και άλλοι στοχαστές σχετικά με τη συνείδηση. Αυτό περιλαμβάνει την ενδελεχή εξέταση των βασικών εννοιών, προϋποθέσεων και λογικών πλαισίων

που παρουσιάζονται στα έργα τους. Στόχος είναι να κατανοηθούν οι βασικές αρχές των αντίστοιχων θεωριών τους και η λογική που τις διέπει.

- Να συγκρίνει και να αντιπαραβάλλει τις προοπτικές των Penrose και Churchland. Αναλύοντας τα έργα τους, η έρευνα αυτή αποσκοπεί στον εντοπισμό των σημείων σύγκλισης και απόκλισης μεταξύ των δύο μελετητών. Θα δοθεί έμφαση στην ανάδειξη των αντιθετικών απόψεων, μεθοδολογιών και παραδοχών που στηρίζουν τις θεωρίες τους.
- Να αξιολογήσει τα δυνατά και αδύνατα σημεία των επιχειρημάτων των δυο θεωρητικών. Αυτό περιλαμβάνει μια κριτική αξιολόγηση της λογικής συνοχής, των εμπειρικών στοιχείων και της επιστημονικής ορθότητας των θεωριών τους. Η έρευνα θα αναλύσει την αποτελεσματικότητα των επιχειρημάτων τους στην αντιμετώπιση του κεντρικού προβλήματος της συνείδησης και τη συμβατότητά τους με τις υπάρχουσες αντιλήψεις και τα επιστημονικά ευρήματα.
- Να εξετάσει τις φιλοσοφικές και επιστημονικές επιπτώσεις των θεωριών των Penrose και Churchland. Η έρευνα έχει ως στόχο να διερευνήσει τις ευρύτερες επιπτώσεις των θεωριών τους στην κατανόηση της συνείδησης, της σχέσης νόησης-εγκεφάλου και της υπόθεσης υπολογισιμότητας της συνείδησης. Θα διερευνήσει πώς οι προοπτικές τους ευθυγραμμίζονται ή αμφισβητούν τα επικρατούντα φιλοσοφικά και επιστημονικά πλαίσια.
- Να συμβάλει στη συνεχιζόμενη συζήτηση για τη φύση της συνείδησης. Αναλύοντας κριτικά τα επιχειρήματα των Penrose και Churchland, η παρούσα έρευνα στοχεύει να προσφέρει ιδέες και να συμβάλει στο υπάρχον σώμα γνώσεων σχετικά με τη συνείδηση. Να εντοπίσει τομείς σύγκλισης ή απόκλισης μεταξύ των θεωριών τους και θα προτείνει πιθανούς δρόμους για περαιτέρω έρευνα και πειραματισμό.

Μέσω αυτών των ερευνητικών στόχων, η παρούσα εργασία επιδιώκει να φωτίσει συγκεκριμένες πλευρές της συζήτησης μεταξύ των δύο στοχαστών αλλά και δύο σημαντικών ρευμάτων σκέψης, βάζοντας το λιθαράκι της σε μια διαδικασία που έρχεται από το μακρινό παρελθόν του πρώιμου φιλοσοφικού στοχασμού, ενώ ανοίγεται διάπλατα σε ένα μέλλον που δε θα σταματήσει να μας εκπλήσσει.

1.4 Μεθοδολογία

Η μεθοδολογία που χρησιμοποιείται στην παρούσα έρευνα περιλαμβάνει μια ανάλυση των επιχειρημάτων που παρουσιάζουν ο Sir Roger Penrose και ο Paul Churchland στα αντίστοιχα έργα τους. Τα ακόλουθα βήματα περιγράφουν την προσέγγιση που χρησιμοποιήθηκε στην παρούσα μελέτη:

- *Ανασκόπηση της βιβλιογραφίας:* Διεξαγωγή εκτενούς ανασκόπησης της σχετικής βιβλιογραφίας για τη φύση της συνείδησης, συμπεριλαμβανομένων των έργων των Penrose, Churchland και άλλων σημαντικών μελετητών. Προσδιορισμός των κύριων θεμάτων, των θεωριών και των συζητήσεων στα σχετικά πεδία, με σκοπό την απόκτηση της επιθυμητής θεωρητικής βάσης για την κατανόηση του πλαισίου των συνεισφορών των Penrose και Churchland.
- *Ανάλυση κειμένων:* Μελέτη των έργων «Shadows of the Mind» και «The Engine of Reason» για τον εντοπισμό των βασικών επιχειρημάτων, των εννοιών και των ισχυρισμών που προβάλλει ο κάθε συγγραφέας. Καταγραφή των βασικών προϋποθέσεων, του θεωρητικού πλαισίου και των εμπειρικών στοιχείων που χρησιμοποιούνται για την υποστήριξη των αντίστοιχων θέσεων.
- *Συγκριτική ανάλυση:* Συστηματική σύγκριση των προοπτικών, θεωριών και μεθοδολογιών του Penrose και του Churchland. Εντοπισμός των σημείων συμφωνίας και διαφωνίας, με έμφαση στις θεμελιώδεις διαφορές των προσεγγίσεών τους για την κατανόηση της συνείδησης. Ανάλυση των επιχειρημάτων κάθε μελετητή και αξιολόγηση της λογικής συνοχής, της εμπειρικής θεμελίωσης και της θεωρητικής συνέπειας των θέσεών τους.
- *Κριτική αξιολόγηση:* Αξιολόγηση των επιχειρημάτων των Penrose και Churchland χρησιμοποιώντας κριτήρια όπως η λογική ορθότητα, η εμπειρική τεκμηρίωση, η επεξηγηματική ισχύς και η συμβατότητα με την υπάρχουσα γνώση. Αναφορά σε κριτικές και εναλλακτικές προοπτικές που διατυπώνονται από άλλους μελετητές. Εκτίμηση των ευρύτερων επιπτώσεων και πιθανών περιορισμών των επιχειρημάτων τους στο πεδίο της φιλοσοφίας του νου.
- *Σύνθεση και συμπέρασμα:* Σύνθεση των ευρημάτων της συγκριτικής ανάλυσης και της κριτικής αξιολόγησης με κατεύθυνση την ανάπτυξη μιας ολοκληρωμένης κατανόησης της αντιπαράθεσης μεταξύ Penrose και Churchland. Σύνοψη των κύριων πλεονεκτημάτων, μειονεκτημάτων και επιπτώσεων των επιχειρημάτων τους. Εξαγωγή τεκμηριωμένων συμπερασμάτων σχετικά με τη φύση της συνείδησης με βάση την ανάλυση των θεωριών τους και το ευρύτερο πλαίσιο της έρευνας της συνείδησης.

Χρησιμοποιώντας αυτή τη μεθοδολογία, η παρούσα έρευνα στοχεύει να παράσχει μια ισορροπημένη κριτική ανάλυση της συζήτησης μεταξύ των Penrose και Churchland, προσφέροντας πληροφορίες για τη φύση της συνείδησης, τις δυνατότητες και τα όρια των αντίστοιχων θέσεών τους, αλλά και τη γενικότερη αντιπαράθεση γύρω από τα ερωτήματα που τέθηκαν στην αρχή του κεφαλαίου.

Κεφάλαιο 2 : Το πρόβλημα νου-σώματος

«Ό,τι δέχτηκα μέχρι πρότινος ως αληθέστατο, το παρέλαβα από τις αισθήσεις ή δια των αισθήσεων. Όμως τούτες τις συνέλαβα να σφάλουν» (Descartes, 2008, p.39).

Η συζήτηση για τη σχέση νου και σώματος κατέχει εξέχουσα θέση στη φιλοσοφία του νου. Εναλλακτικές θεωρίες έχουν αναπτυχθεί για να προσφέρουν λύσεις στα σχετικά ζητήματα που εγείρονται, ζητήματα που παραμένουν εν πολλοίς άλυτα. Το κύριο πρόβλημα οφειλόταν στη θεώρηση ότι, εφόσον το σώμα είναι φυσικό και ο νους μη φυσικός, το σώμα εκτείνεται, καταλαμβάνει κάποιο χώρο και ο νους δεν καταλαμβάνει κανένα χώρο, το σώμα είναι ένας τύπος ουσίας και ο νους είναι ένας άλλος τύπος ουσίας, πώς μπορούν να συσχετιστούν; Παρακάτω εξετάζονται τόσο ορισμένες φιλοσοφικές θέσεις που προτάθηκαν από διακεκριμένους στοχαστές του 17^{ου} αιώνα όσο και πιο σύγχρονες θεωρήσεις που διέπονται φυσικά από μεγαλύτερη συνέπεια και επιστημονική συστηματικότητα.

2.1 Ο δυϊσμός του Descartes

«Αλλά πείστηκα ότι δεν υπάρχει τίποτε απολύτως στον κόσμο, ούτε ουρανός, ούτε γη, ούτε πνεύματα, ούτε σώματα. Μήπως λοιπόν δεν υπάρχω ούτε εγώ; Όχι, αν πείστηκα για κάτι, τότε είναι βέβαιο ότι υπάρχω. Αλλά υπάρχει κάποιος μυστηριώδης απατεώνας, παντοδύναμος και παμπόνηρος, που με ξεγελά με πανουργία διαρκώς. Αν όμως με ξεγελά, είναι αναμφίβολο ότι υπάρχω. Κι ας με ξεγελά όσο θέλει, αφού δε θα με κάνει ποτέ να μην είμαι τίποτα όσο θα σκέφτομαι ότι είμαι κάτι. Ωστε μετά από επαρκές ζύγισμα των πάντων, πρέπει εντέλει να καταλήξω ότι η απόφαση *Εγώ είμαι, εγώ υπάρχω* αληθεύει αναγκαία όποτε την προφέρω ή τη συλλαμβάνω στο πνεύμα μου» (Descartes, 2008, p.44).

Ο Γάλλος φιλόσοφος και μαθηματικός Renes Descartes ανέπτυξε εκτενώς την άποψή του για τη σχέση μεταξύ νου και σώματος στο έργο του «Στοχασμοί περί της Πρώτης Φιλοσοφίας» (*Meditationes de prima philosophia*) το 1641. Εκεί περιγράφει την ψυχή σαν «άυλο, σκεπτόμενο πράγμα», που αποτελεί την ουσία του ανθρώπου και εμπεριέχει όλες τις σκέψεις, ελπίδες και αμφιβολίες του, καθώς και την πίστη του. Κατά τον Καρτέσιο, σώμα και ψυχή βρίσκονται σε αλληλεπίδραση μεταξύ τους, η ψυχή όμως μπορεί να υπάρξει και ανεξάρτητα από το σώμα.

«Αλλά τι είμαι λοιπόν; Σκεπτόμενο πράγμα (*Res cogitans*). Τι είναι αυτό; Είναι ένα πράγμα που αμφιβάλλει, νοεί, βεβαιώνει, αρνείται, θέλει, δεν θέλει, φαντάζεται επίσης αισθάνεται» (Descartes, 2008, p.46).

Αναζητώντας μια βεβαιότητα που να αποτελεί απόδειξη ενάντια σε κάθε αμφιβολία, ο Descartes ξεκινά τους *Στοχασμούς* με σκεπτικιστικά επιχειρήματα που αποσκοπούν να δείξουν τι μπορεί να τεθεί υπό αμφισβήτηση και συνεχίζει ανακαλύπτοντας τη φύση του ανθρώπινου νου και πώς αυτός είναι καλύτερα γνωστός από το σώμα. Έχοντας θέσει υπό αμφισβήτηση την ίδια την ύπαρξη του οικείου υλικού κόσμου, βρίσκει βεβαιότητα στη γνώση του περιεχομένου του δικού του νου και καταλήγει στο συμπέρασμα ότι ο ίδιος ως

νους δεν μπορεί να είναι μέρος του υλικού κόσμου. Μια νέα μορφή σκεπτικισμού, ο σκεπτικισμός για τον εξωτερικό κόσμο, και μια νέα αντίληψη του νου, ο καρτεσιανός νους, προκύπτουν από την καρτεσιανή αναζήτηση της βεβαιότητας. Ο καρτεσιανός νους διαχωρίζεται από τον υλικό κόσμο τόσο επιστημολογικά όσο και μεταφυσικά, με συνέπεια αυτός ο διαχωρισμός να δημιουργεί ξεχωριστά προβλήματα: το επιστημολογικό πρόβλημα της επανασύνδεσης του νου με τον κόσμο - το πρόβλημα του εξωτερικού κόσμου - και το μεταφυσικό πρόβλημα της επανασύνδεσης του νου με το σώμα - το πρόβλημα νους-σώμα.

«Πόσες φορές έχω ονειρευτεί ότι βρισκόμουν σε αυτές τις οικείες συνθήκες, ότι ήμουν ντυμένος και καταλάμβανα αυτή τη θέση δίπλα στη φωτιά, ενώ ήμουν ξαπλωμένος, χωρίς ρούχα, στο κρεβάτι μου; Αυτή τη στιγμή...κοιτάζω αυτό το χαρτί, με τα μάτια ορθάνοιχτα... αλλά δεν μπορώ να ξεχάσω ότι και σε άλλες στιγμές έχω εξαπατηθεί στον ύπνο μου από παρόμοιες ψευδαισθήσεις· και εξετάζοντας λεπτομερώς αυτές τις περιπτώσεις, αντιλαμβάνομαι τόσο ξεκάθαρα, ότι δεν υπάρχουν σημάδια για να διακρίνουμε μετά βεβαιότητας την κατάσταση της εγρήγορσης από αυτή του ύπνου, ώστε μένω έκπληκτος...»(Descartes, 2008, p.40).

Μπορούμε να αναλύσουμε την άποψη του Descartes σε δυο μέρη· το ένα αναφέρεται στη διάκριση μεταξύ ουσιών και το άλλο στη διάκριση μεταξύ ιδιοτήτων. Ένας τρόπος να κατανοήσουμε αυτή τη διάκριση είναι μέσω της διάκρισης μεταξύ ονομάτων και κατηγορημάτων. Αντίστοιχη με τη διάκριση μεταξύ ουσιών και κατηγορημάτων είναι η διάκριση μεταξύ δύο ειδών δυϊσμού. Όπως για τον δυϊστή των ιδιοτήτων οι νοητικές ιδιότητες - όπως το να αισθάνεσαι έναν πόνο ή να σκέφτεσαι ένα φαγητό - δεν ταυτίζονται με καμία φυσική ιδιότητα, έτσι και για το δυϊστή της ουσίας υπάρχουν νοητικές ουσίες - ο νους - οι οποίες είναι δεν ταυτίζονται με κανένα φυσικό πράγμα. Ο Descartes μοιραζόταν και τις δυο αυτές αντιλήψεις.

Μια άλλη διάκριση εντός του δυϊσμού αξίζει να γίνει εδώ. Για τον Descartes, ο νους μπορεί να επιδρά αιτιωδώς στα σώματα, αλλά και αντιστρόφως, οι νοητικές ουσίες είναι ικανές να επιφέρουν φυσικά αποτελέσματα. Έτσι, παρά το γεγονός ότι συνιστούν διαφορετικά είδη πραγμάτων, νους και σώμα μπορούν να αλληλεπιδρούν. Αυτό το ρεύμα σκέψης αποκαλείται *αλληλεπιδραστικός δυϊσμός* (interactive dualism).

Ο δυϊσμός νου-σώματος ονομάζεται μερικές φορές δυϊσμός των ουσιών για να διακριθεί από τον δυϊσμό των ιδιοτήτων· τη θέση ότι οι νοητικές ιδιότητες (όπως το να αισθάνεσαι πόνο, να σκέφτεσαι ένα πρόσωπο) είναι κατά κάποιο τρόπο σημαντικά διαφορετικές ή ανεξάρτητες από τις φυσικές ιδιότητες (όπως το να πυροδοτούνται οι νευρώνες στον εγκέφαλο κάποιου με ένα συγκεκριμένο μοτίβο). Ο δυϊσμός των ιδιοτήτων έχει σκοπό να καταστήσει δυνατό αυτό που αποκαλείται συχνά *θεωρία διπλής όψης* (dual-aspect theory): τα άτομα είναι υλικά αντικείμενα με μια μη φυσική, νοητική διάσταση αλλά χωρίς μη φυσική μορφή, δηλαδή χωρίς άυλη ψυχή.

Ένα ερώτημα που προκύπτει σε αυτό το σημείο είναι: ποια είναι η σχέση μεταξύ του δυϊσμού της ουσίας και του δυϊσμού των ιδιοτήτων; Αν ο δυϊσμός της ουσίας είναι αληθής, προκύπτει λογικά ότι ο δυϊσμός των ιδιοτήτων πρέπει να είναι επίσης αληθής; Μήπως το αντίστροφο;

Ορισμένοι επικριτές δεν αποδέχθηκαν τη θεωρία του Descartes για τον εξής λόγο: πώς μπορεί μια φυσική ουσία να επικοινωνήσει, να σχετιστεί, να έρθει σε επαφή με τη νοητική ουσία; Είναι δυνατόν να εξηγηθεί το φυσικό με το νοητικό και το νοητικό με το φυσικό, δεδομένου ότι ο νους είναι ένα είδος ουσίας και το σώμα είναι ένα άλλο είδος ουσίας; Εάν η επίφυση μπορεί να ρυθμίσει αυτά τα δύο, μπορούμε να πούμε ότι η επίφυση είναι φυσική; Αν είναι έτσι, τότε, εδώ μια φυσική ουσία έρχεται σε επαφή με τη νοητική ουσία. Προφανώς, δεν υπάρχει αμφιβολία ότι ο αδένας είναι φυσικός. Επομένως, η δυνατότητα επαφής μέσω της επίφυσης (η οποία είναι φυσική) δεν μπορεί να γίνει αποδεκτή.

Ο Spinoza, λίγα χρόνια αργότερα στη μνημειώδη *Ηθική* του, θα απορρίψει τον καρτεσιανό δυϊσμό. Η θεμελιώδης καρτεσιανή διάκριση μεταξύ της ύλης ή των σωμάτων, των οποίων η βασική ιδιότητα είναι ότι εκτείνονται χωρικά, και του νου, του οποίου η βασική ιδιότητα είναι ότι σκέφτεται, αντικρούεται από τη μονιστική προσέγγιση του Spinoza. Για τον Spinoza υπάρχει μόνο μία ουσία, καθώς αν υπήρχαν περισσότερες από μία, θα περιόριζαν η μία την άλλη και θα κατέστρεφαν την αμοιβαία ουσιαστικότητά τους.

2.2 Ο μονισμός του Spinoza

Ο Ολλανδός φιλόσοφος και λειαντής φακών Baruch Spinoza (1632-1677) κατηγορήθηκε και αγαπήθηκε όσοι λίγοι φιλόσοφοι για τις -προκλητικές για την εποχή του- θέσεις του. Σύμφωνα με τον Spinoza, υπάρχει μία και μοναδική αυθύπαρκτη οντότητα, η *υπόσταση*, η οποία λέγεται και Θεός ή Φύση. Αυτή η μοναδική υπόσταση διαθέτει άπειρα κατηγορήματα, καθένα από τα οποία εκφράζει την αιώνια και άπειρη ουσία του Θεού. Τα κατηγορήματα είναι «αυτό που η διάνοια (ο νους) αντιλαμβάνεται για την υπόσταση, ως αυτό που συγκροτεί την ουσία της» (Curley, 1994, p.121).

Τα ανθρώπινα όντα μπορούν να γνωρίζουν μόνο δύο κατηγορήματα, δηλαδή το νοητικό (σκέψη) και το φυσικό (έκταση). Ο Spinoza θεωρεί ότι ο Θεός ή Φύση είναι ένας συνδυασμός του νου και του σώματος. Επομένως, όπου υπάρχει χώρος ή ύλη, υπάρχει ψυχή ή νους και αντίστροφα. Τα δυο κατηγορήματα (σκέψη και έκταση), όντας ουσιώδη για τη φύση της υπόστασης, πρέπει να υπάρχουν οπουδήποτε βρίσκεται η υπόσταση, δηλαδή παντού. Η σκέψη και η έκταση είναι το καθένα άπειρο στο είδος του, αλλά όχι απόλυτα άπειρο, δηλαδή ούτε η σκέψη ούτε η έκταση είναι το μοναδικό κατηγορήμα· αφού υπάρχουν πολλά άλλα κατηγορήματα του Θεού, κανένα από αυτά, δεν μπορεί να χαρακτηριστεί απόλυτα άπειρο. Αυτά τα κατηγορήματα χαρακτηρίζονται από αντιληπτική ανεξαρτησία, καθένα τους συλλαμβάνεται δηλαδή δια του εαυτού του, χωρίς αναφορά σε άλλα κατηγορήματα. Ακόμα, τα κατηγορήματα της υπόστασης δεν μπορούν να αλληλεπιδράσουν αιτιακά το ένα στο άλλο, βρίσκονται δηλαδή σε σχέση αιτιακής ανεξαρτησίας. Ο Spinoza θεωρεί ότι το φυσικό δεν μπορεί να εξηγηθεί από το νοητικό και το νοητικό δεν μπορεί να εξηγηθεί από το φυσικό. Τα δύο είναι διαφορετικά μεταξύ τους όπως δύο παράλληλες γραμμές των ραγών του τρένου, που δεν συναντιούνται ποτέ. Δεν μπορεί να υπάρξει αιτιώδης σχέση μεταξύ του νοητικού και του φυσικού.

Οι *τρόποι* της υπόστασης, αποτελούν το δεύτερο είδος οντοτήτων της σπινοζικής οντολογίας. Οι τρόποι είναι τα ενικά πράγματα που αντιλαμβανόμαστε στην καθημερινή μας ζωή λόγου χάρη τα ζώα, οι άνθρωποι, οι ιδέες, τα φυσικά αντικείμενα κλπ.

«Με τον όρο *τρόπο* καταλαβαίνω τις διαθέσεις (affections) μιας ουσίας, ή αυτό το οποίο περιέχεται σε μια άλλη, μέσω της οποίας επίσης συλλαμβάνεται» (Curley, 1994, p.121).

Ενώ η *υπόσταση* είναι η πιο βασική και ουσιαστική οντότητα, και ο *τρόπος* είναι μια ιδιαίτερη πτυχή αυτής της ύπαρξης, τα *κατηγορήματα* είναι τρόποι ύπαρξης.

Για να δεχτούμε τη μεταφυσική θεωρία που προτείνει ο Spinoza, πρέπει να κατανοήσουμε καλύτερα ορισμένες από τις λέξεις που χρησιμοποιεί συχνά. Ο όρος "υπόσταση" που χρησιμοποιεί ο Spinoza είναι παρόμοιος με τον όρο "ουσία" που χρησιμοποιεί ο Descartes και έχει κληρονομηθεί από τον Αριστοτέλη και τον μεσαιωνικό σχολαστικισμό:

«Με τον όρο υπόσταση εννοώ αυτό που είναι στον εαυτό του και αυτό που γίνεται κατανοητό μόνο από τον εαυτό του, δηλαδή αυτό που δεν χρειάζεται άλλον όρο για ένα άλλο πράγμα, από το οποίο πρέπει να συγκροτηθεί» (Curley, 1994, p.121).

Ορίζοντας έτσι τον όρο "υπόσταση", ο Spinoza λέει ότι ο μόνος υποψήφιος για μια τέτοια ουσία είναι η πραγματικότητα στο σύνολό της. Αυτό συμβαίνει για δύο λόγους. Πρώτον, αυτό που υπάρχει, που είναι, είναι. Δεύτερον, αυτό που είναι, είναι όλα όσα είναι· επομένως, δεν μπορεί να υπάρχει τίποτα από το οποίο να εξαρτάται η ύπαρξη του συνόλου της ύπαρξης. Αν υπήρχε μια αιτία, η μια ουσία είναι η αιτία για τον εαυτό της, λέει ο Spinoza.

Ο Descartes αναγνώρισε τον Θεό ως το μόνο ον που μπορεί να κατανοηθεί ως μη εξαρτώμενο από κανένα άλλο πράγμα. Ωστόσο, ο Descartes είπε επίσης ότι τόσο ο νους όσο και το σώμα είναι ουσίες που «μπορούν να υπάρξουν μόνο με τη βοήθεια της συνδρομής του Θεού».

Ο Descartes καταλήγει σε αυτόν τον ορισμό δηλώνοντας:

«Στην περίπτωση των δημιουργημένων πραγμάτων, ορισμένα είναι τέτοιας φύσης που δεν μπορούν να υπάρξουν χωρίς άλλα πράγματα, ενώ ορισμένα χρειάζονται μόνο τη συνηθισμένη συμβολή του Θεού για να υπάρξουν. Κάνουμε αυτή τη διάκριση ονομάζοντας τα τελευταίες *ουσίες* και τα πρώτες *ιδιότητες* ή *κατηγορήματα* αυτών των ουσιών» (Miller V., Miller R., p.47).

Ενώ ο Spinoza αποδέχεται τη βασική αρχή του καρτεσιανού ορισμού της ουσίας, δεν υιοθετεί την άποψη του Ντεκάρτ, ότι ο νους και το σώμα συνιστούν ουσίες. Όταν ο Spinoza λέει ότι η ουσία (ή υπόσταση) δεν απαιτεί την αντίληψη οποιουδήποτε άλλου πράγματος πέρα από τον εαυτό της, αποκλείει το νου και το σώμα ως ουσίες, αφού η ουσία τους προϋποθέτει κάτι άλλο, δηλαδή τον Θεό. Η ουσία πρέπει να είναι κάτι που δεν περιορίζεται ή ορίζεται από κάτι άλλο. Η ουσία δεν μπορεί να είναι μια ποιότητα ή μια σχέση, καθώς

αυτοί οι όροι υπονοούν εξάρτηση από πράγματα που διαφέρουν από τον εαυτό τους. Όλες οι σχέσεις και οι ποιότητες είναι δευτερεύουσες και εξαρτώνται από την ουσία.

Η *παράλληλία* (parallelism) του Spinoza δεν συνίσταται μόνο στην άρνηση οποιασδήποτε σχέσης πραγματικής αιτιότητας ανάμεσα στο πνεύμα και το σώμα, αλλά απαγορεύει και την οποιαδήποτε υπεροχή του ενός έναντι του άλλου. Η άρνηση της ανωτερότητας της ψυχής, της προνομιακής θέσης των νοητικών καταστάσεων έναντι του σώματος εκ μέρους του Spinoza, δε στοχεύει στην εγκαθίδρυση του αντίστροφου είδους ανωτερότητας, του σώματος έναντι του νου (Deleuze, 1981).

«Ό,τι είναι ενέργημα στην ψυχή είναι κατ'αναγκαιότητα ενέργημα και στο σώμα, ό,τι είναι πάθος στο σώμα είναι κατ'αναγκαιότητα πάθος στην ψυχή» (Deleuze, 1981, p.33)

2.4 Θεωρία Ταυτότητας Νου-Εγκεφάλου

Η θεωρία της ταυτότητας νου-εγκεφάλου (Mind-Brain Identity Theory) αποτελεί κομμάτι της υλιστικής φιλοσοφικής παράδοσης που αναπτύχθηκε κυρίως στα μέσα του 20^{ου} αιώνα με κυριότερους εκπροσώπους τον J.J.C. Smart και U.T. Place. Στη βιβλιογραφία εμφανίζεται συχνά και με τους όρους φυσικαλισμός (physicalism) ή εξαιρετικισμός υλισμός (reductive materialism). Η θεωρία βασίζεται στην παραδοχή ότι ο νους και ο εγκέφαλος είναι το ίδιο πράγμα. Ο νους αποτελεί μέρος του φυσικού σώματος και η κατάσταση του νου είναι ίδια με τις εγκεφαλικές διεργασίες· η νοητική κατάσταση ταυτίζεται με τη φυσική κατάσταση του εγκεφάλου.

Ο Βρετανός φιλόσοφος και ψυχολόγος U.T. Place, θέτει ορισμένα θεμέλια της θεωρίας στο έργο του «Είναι η συνείδηση εγκεφαλική διεργασία;» (Is Consciousness a Brain Process?, 1954). «Η άποψη ότι υπάρχει μια ξεχωριστή κατηγορία συμβάντων, τα νοητικά συμβάντα, τα οποία δεν μπορούν να περιγραφούν με όρους των εννοιών που χρησιμοποιούν οι φυσικές επιστήμες, δεν διαθέτει πλέον την καθολική και αδιαμφισβήτητη αποδοχή μεταξύ των φιλοσόφων και των ψυχολόγων που είχε κάποτε» (Place, 1956, p.1). Για τον Place, η αποδοχή εσωτερικών διεργασιών δεν έχει ως αναγκαία λογική συνέπεια τον δυϊσμό.

Αναζητώντας μια εναλλακτική λύση στην κλασική δυϊστική θέση, σύμφωνα με την οποία οι νοητικές καταστάσεις διαφέρουν σε οντολογικό επίπεδο από τις φυσιολογικές καταστάσεις με τις οποίες θεωρείται ότι συσχετίζονται, ο Place υποστήριξε ότι οι αισθήσεις και αντίστοιχα νοητικά συμβάντα θα μπορούσαν να είναι διεργασίες στον εγκέφαλο. Μέσα από παραδείγματα, ο Place ξεκαθαρίζει ότι ενώ στις γνωσιακές δηλώσεις (προτάσεις που εισάγονται με γνωσιακές έννοιες, λ.χ., «γνωρίζω», «πιστεύω», «κατανοώ», «θυμάμαι») μια κατά βάση συμπεριφοριστική ανάλυση είναι δυνατή, μια αντίστοιχη αναγωγή δηλώσεων σχετικών με τις καταστάσεις της συνείδησης σε δηλώσεις εγκεφαλικών διεργασιών αποτυγχάνει. Μπορούμε, για παράδειγμα, να μιλήσουμε για τις νοητικές μας εικόνες χωρίς να γνωρίζουμε τι συμβαίνει μέσα στον εγκέφαλό μας. Ωστόσο, για τον Place, η δήλωση «η συνείδηση είναι εγκεφαλική διεργασία», ως εύλογη επιστημονική υπόθεση, δεν είναι

κατ'ανάγκη αληθής ούτε κατ'ανάγκη ψευδής, όμοια με τη δήλωση «η αστραπή είναι μια κίνηση ηλεκτρικών φορτίων» (Place, 1956, p.2).

Ωστόσο το πρόβλημα της εξήγησης των ενδοσκοπικών αναφορών με όρους εγκεφαλικών διεργασιών παρέμενε, δεδομένου ότι αυτές οι δηλώσεις συνήθως αναφέρονται σε οντότητες που δεν ταιριάζουν με τη φυσικαλιστική εικόνα. Για να λύσει αυτό το πρόβλημα, ο Place περιγράφει ένα λογικό σφάλμα που διαπράττουν συνήθως φυσιολόγοι και ψυχολόγοι, το οποίο ονομάζει *φαινομενολογικό σφάλμα*. Αυτός ο όρος αναφέρεται στη λανθασμένη υπόθεση ότι οι ενδοσκοπικές παρατηρήσεις κάποιου δηλώνουν «την πραγματική κατάσταση των πραγμάτων σε κάποιο μυστηριώδες εσωτερικό περιβάλλον». Συγκεκριμένα, το εν λόγω σφάλμα, για τον Place, «έγκειται στην υπόθεση πως όταν το υποκείμενο περιγράφει την εμπειρία του, όταν περιγράφει το πώς φαίνονται και πώς ηχούν τα πράγματα, το τι οσμή και τι γεύση έχουν, το πώς τα αισθάνεται, περιγράφει κυριολεκτικά τις ιδιότητες των αντικειμένων και των συμβάντων σε ένα ιδιαίτερο είδος εσωτερικής κινηματογραφικής ή τηλεοπτικής οθόνης, ό,τι συνήθως ονομάζεται στη σύγχρονη ψυχολογική γραμματεία *φαινόμενο πεδίο*» (Place, 1956, p.6).

Όπως αναφέρει ο Place, «όταν περιγράφουμε το μετείκασμα ως πράσινο, δεν λέμε ότι υπάρχει κάτι, το μετείκασμα, που είναι πράσινο· λέμε ότι έχουμε το είδος εμπειρίας που υπό κανονικές συνθήκες έχουμε, όταν κοιτάζουμε μια πράσινη φωτεινή κηλίδα, και έχουμε μάθει να περιγράφουμε σαν τέτοια» (Place, 1956, p.7). Για να δοθεί μια σαφής εξήγηση στην ενδοσκοπική παρατήρηση του υποκειμένου, σύμφωνα με τον Place, αρκεί να δειχθεί ότι η εγκεφαλική διεργασία που εκτελεί το υποκείμενο με σκοπό να περιγράψει την εμπειρία του, με αυτόν τον ιδιαίτερο τρόπο, είναι το είδος της διεργασίας που συμβαίνει κανονικά όταν υπάρχει πράγματι κάτι στο περιβάλλον που αντιστοιχεί στην περιγραφή του.

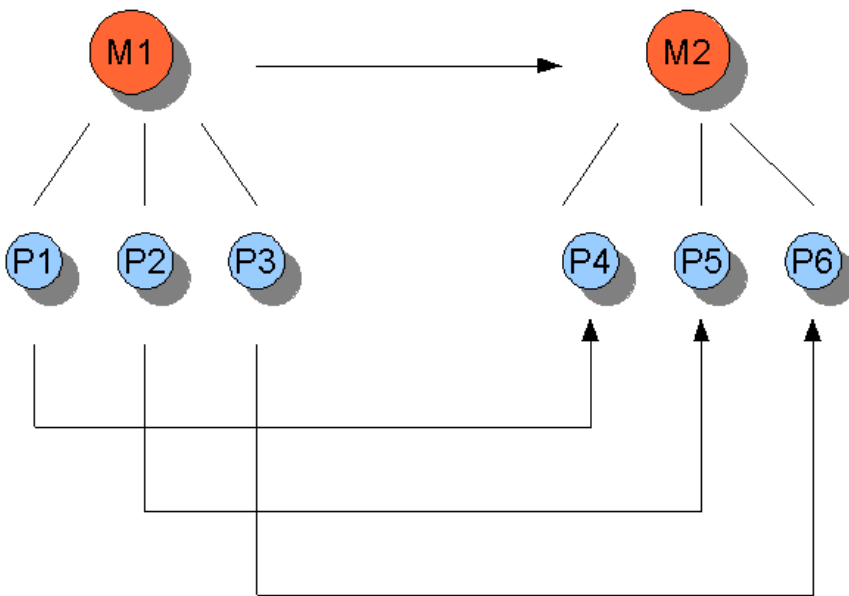
2.5 Λειτουργισμός

Ο λειτουργισμός (functionalism) σαν όρος στη φιλοσοφία του νου εισήχθη στα μέσα του 20^{ου} αιώνα και συνιστά μια οικογένεια θεωριών που εξετάζει τη φύση των νοητικών καταστάσεων. Αναπτύχθηκε σε αντιπαράθεση με τις επικρατούσες απόψεις που κυριαρχούσαν τότε, τη θεωρία ταυτότητας και το συμπεριφορισμό. Σύμφωνα με τον λειτουργισμό, οι νοητικές καταστάσεις προσδιορίζονται από αυτό που επιτελούν και όχι από αυτό από το οποίο αποτελούνται. Αυτό που κάνει κάτι σκέψη, επιθυμία, πόνο (ή οποιοδήποτε άλλο είδος νοητικής κατάστασης) δεν εξαρτάται από την εσωτερική του σύσταση, αλλά αποκλειστικά από τη λειτουργία του, το ρόλο που παίζει μέσα στο γνωστικό σύστημα του οποίου αποτελεί μέρος. Πιο συγκεκριμένα, για τις λειτουργικές θεωρίες, η ταυτότητα μιας νοητικής κατάστασης καθορίζεται από τις αιτιώδεις σχέσεις της με τα αισθητηριακά ερεθίσματα, τις υπόλοιπες νοητικές καταστάσεις και τη συμπεριφορά του υποκειμένου. Αφετηρία της διδασκαλίας του λειτουργισμού θεωρείται η σύγκριση των νοητικών λειτουργιών με των λειτουργιών ενός υπολογιστικού συστήματος.

Ερχόμενος σε αντιπαράθεση με τον συμπεριφορισμό, ο λειτουργισμός διατήρησε την παραδοσιακή ιδέα ότι οι νοητικές καταστάσεις είναι εσωτερικές καταστάσεις σκεπτόμενων πλασμάτων. Ο λειτουργισμός υπερβαίνει τον συμπεριφορισμό επιτρέποντας να ορίζονται

οι νοητικές καταστάσεις όχι μόνο με όρους συμπεριφοράς (είσοδοι και έξοδοι) αλλά και με όρους άλλων νοητικών καταστάσεων. Σημαντική θεωρείται και η συνεισφορά του Αμερικανού φιλοσόφου Hilary Putnam στην ανάπτυξη του λειτουργισμού, μέσω της κριτικής που άσκησε στη βασική υπόθεση των θεωριών ταυτότητας.

Στο έργο του με τίτλο «Η φύση των νοητικών καταστάσεων» (The Nature of Mental States), ένας από τους κυριότερους εκπροσώπους του λειτουργισμού, ο Putnam διατύπωσε, ίσως τη σοβαρότερη ως τότε, αντίρρηση στις θεωρίες της ταυτότητας νου-εγκεφάλου. Το επιχείρημα του Putnam, γνωστό και ως *δυνατότητα πολλαπλής πραγμάτωσης* (multiple realizability) μπορεί να συνοψιστεί ως εξής: Σύμφωνα με τη θεωρία ταυτότητας, σε κάθε νοητική κατάσταση αντιστοιχεί μια μοναδική φυσική-χημική κατάσταση του εγκεφάλου. Με άλλα λόγια, μια μορφή ζωής να μπορεί να βρίσκεται σε αυτή τη νοητική κατάσταση εάν και μόνο εάν βρίσκεται σε αυτή τη φυσική κατάσταση. Εύλογη είναι ωστόσο και η εμπειρική υπόθεση, ότι διαφορετικές μορφές ζωής μπορούν να βρίσκονται στην ίδια νοητική κατάσταση, χωρίς να συμβαίνουν τα ίδια πράγματα στον εγκέφαλό τους. Η προηγούμενη συλλογιστική, φανερώνει την αδυναμία των υποθέσεων στις οποίες βασίζονται οι θεωρίες ταυτότητας. Η παραπάνω έννοια προκαλεί την ιδέα μιας αυστηρής, ένα προς ένα, αντιστοιχίας μεταξύ ψυχικών καταστάσεων και συγκεκριμένων φυσικών καταστάσεων. Υπονοεί ότι οι νοητικές διαδικασίες δεν είναι συνδεδεμένες με ένα μοναδικό φυσικό υπόστρωμα, αλλά μπορούν να υλοποιηθούν με ποικίλους φυσικούς μηχανισμούς (Putnam, The Nature of Mental States, 2010).



Εικόνα 1: Μια απεικόνιση της δυνατότητας πολλαπλής πραγμάτωσης. Το *M* συμβολίζει το νοητικό και το *P* το φυσικό. Παρατηρείται ότι περισσότερα από ένα *P* μπορούν να υλοποιήσουν ένα *M*, αλλά όχι το αντίστροφο. Οι αιτιώδεις σχέσεις μεταξύ καταστάσεων αναπαρίστανται με τα βέλη (το *M1* πηγαίνει στο *M2* κ.λπ.).

Ο Putnam προσφέρει ένα ακόμα επιχείρημα υπέρ της πολλαπλής δυνατότητας πραγμάτωσης, εισάγοντας την έννοια του *λειτουργικού ισομορφισμού*. Η έννοια

χαρακτηρίζεται ως εξής: «Δύο συστήματα είναι λειτουργικά ισομορφικά εάν υπάρχει μια αντιστοιχία μεταξύ των καταστάσεων του ενός και των καταστάσεων του άλλου που διατηρεί τις λειτουργικές σχέσεις» (Putnam, 2010, *Philosophy and our Mental Life*, p.1). Στην περίπτωση των υπολογιστών, δύο μηχανές είναι λειτουργικά ισομορφικές εάν και μόνο εάν οι διαδοχικές σχέσεις μεταξύ των καταστάσεων στην πρώτη αντικατοπτρίζουν ακριβώς τις διαδοχικές σχέσεις μεταξύ των καταστάσεων στην άλλη. Επομένως, ένας υπολογιστής από τσιπ πυριτίου και ένας υπολογιστής από γρανάζια και τροχούς μπορεί να είναι λειτουργικά ισομορφικοί αλλά δομικά διαφορετικοί. Ο λειτουργικός ισομορφισμός συνεπάγεται τη δυνατότητα πολλαπλής πραγμάτωσης.

2.6 Η κριτική του Gilbert Ryle στο καρτεσιανό δόγμα

Σύμφωνα με τον Καρτέσιο, τα ανθρώπινα σώματα είναι στο χώρο και υπόκεινται στους μηχανικούς νόμους που διέπουν όλα τα άλλα σώματα στο χώρο, ενώ οι σωματικές διεργασίες και καταστάσεις είναι ορατές και εποπτεύσιμες από εξωτερικούς παρατηρητές.

Ο νους, αντίθετα, δεν βρίσκεται στο χώρο, ούτε οι λειτουργίες του υπόκεινται στους μηχανικούς νόμους· η πορεία του νου είναι ιδιωτική. Μόνο εγώ έχω άμεση πρόσβαση στις καταστάσεις και τις διεργασίες του νου μου. Ως εκ τούτου, ένα άτομο βιώνει δύο παράπλευρες ιστορίες: η πρώτη ιστορία είναι δημόσια και αφορά το τι συμβαίνει στο σώμα ενώ, η δεύτερη ιδιωτική και αναφέρεται στο τι συμβαίνει στο νου. Τα συμβάντα στην πρώτη ιστορία είναι συμβάντα στο φυσικό κόσμο, ενώ στη δεύτερη είναι συμβάντα στο νοητικό κόσμο (Ryle, 1949, p.1).

Πώς αλληλοεπηρεάζονται όμως αυτές οι δυο οντότητες; Ό,τι σκέφτεται ο νους, τα πόδια, τα χέρια και η γλώσσα το εκτελούν· ό,τι επιδρά στο αυτί και στο μάτι σχετίζεται κάπως με το τι ο νους αντιλαμβάνεται· οι μορφασμοί και το χαμόγελο προδίδουν τις διαθέσεις του νου· η σωματική τιμωρία ελπίζεται πως οδηγεί σε ηθική βελτίωση. Όμως το πραγματικό «δούναι και λαβείν» ανάμεσα στα επεισόδια της ιδιωτικής ιστορίας και σε εκείνα της δημόσιας ιστορίας παραμένει μυστηριώδες, αφού εξ ορισμού αυτή η συναλλαγή δεν είναι δυνατόν να ανήκει ούτε στη μία ούτε στην άλλη σειρά επεισοδίων. Δεν θα μπορούσε να καταχωριστεί ανάμεσα στα συμβάντα όσα περιγράφονται στην αυτοβιογραφική έκθεση του εσωτερικού βίου ενός ατόμου, ούτε όμως και ανάμεσα στα συμβάντα όσα περιγράφει ένας τρίτος, ένας βιογράφος όταν εξιστορεί την εμφανή, δημόσια πορεία του ατόμου. Η αλληλεπίδραση ανάμεσα στα επεισόδια της ιδιωτικής και της δημόσιας ιστορίας δεν μπορεί να εξεταστεί ούτε με ενδοσκόπηση ούτε με εργαστηριακά πειράματα. Πρόκειται για ένα θεωρητικό μπαλάκι που αενάως στέλνει ο φυσιολόγος στον ψυχολόγο και ο ψυχολόγος επιστρέφει στον φυσιολόγο (Ryle, 1949, p.2).

Η παραπάνω διχοτόμηση έχει τις ρίζες της σε μια βαθύτερη φιλοσοφική παραδοχή που κάνει λόγο για δύο διαφορετικά είδη ύπαρξης ή υπόστασης. Ό,τι υπάρχει ή συμβαίνει μπορεί να έχει είτε την υπόσταση φυσικής ύπαρξης ή είτε αυτή της νοητικής ύπαρξης. Αναγκαίο χαρακτηριστικό των πραγμάτων που έχουν φυσική ύπαρξη είναι ότι βρίσκονται στο χώρο και στο χρόνο· αναγκαίο χαρακτηριστικό των πραγμάτων που έχουν νοητική ύπαρξη είναι ότι βρίσκονται στο χρόνο αλλά όχι στο χώρο. Ό,τι έχει φυσική ύπαρξη

συνίσταται από ύλη ή είναι λειτουργία της ύλης· ό,τι έχει νοητική ύπαρξη συνίσταται από τη συνείδηση ή είναι λειτουργία της συνείδησης (Ryle, 1949, p.2).

Ο Ryle αναφέρεται στο καρτεσιανό δόγμα με τη διάσημη φράση του για το «φάντασμα εντός της μηχανής». Για τον Ryle η θεωρία δεν υποπίπτει απλά σε λάθη λεπτομερειών, αλλά είναι θεμελιωδώς εσφαλμένη, καθώς βασίζεται σε σφάλμα κατηγορίας. Παρουσιάζει τα γεγονότα του νοητικού βίου σαν να ανήκαν σε ένα λογικό τύπο ή κατηγορία, ενώ στην πραγματικότητα ανήκουν σε άλλον. Αρνείται εξ ολοκλήρου την εν λόγω θεωρία αποκαλώντας τη μύθο και επιχειρεί να εξηγήσει μέσα από συγκεκριμένα παραδείγματα τον όρο κατηγοριακά σφάλματα (category errors) (Ryle, 1949, p.4).

Στο πρώτο από αυτά, ένας ξένος βρίσκεται για πρώτη φορά στο πανεπιστήμιο της Οξφόρδης. Αφού ξεναγηθεί στα διάφορα κολέγια, βιβλιοθήκες, γήπεδα, μουσεία, επιστημονικά τμήματα, γραφεία διοικητικών υπηρεσιών, ο επισκέπτης ρωτά «και το Πανεπιστήμιο; Πού είναι το Πανεπιστήμιο; Είδα πού ζουν τα μέλη των κολεγίων, πού εργάζεται ο γραμματέας του Πανεπιστημίου, πού διεξάγουν τα πειράματα τους οι επιστήμονες, και τα λοιπά. Όμως ακόμα να δω το Πανεπιστήμιο όπου ζουν και εργάζονται τα μέλη του Πανεπιστημίου σας.» Το πανεπιστήμιο δεν πρόκειται για ένα ακόμα ίδρυμα δίπλα στα τμήματα, στα γραφεία κλπ που ξεναγήθηκε, αλλά ένας τρόπος με τον οποίο όλα αυτά είναι οργανωμένα. Το σφάλμα του επισκέπτη έγκειται στην αθώα υπόθεση του ότι ήταν σωστό να γίνεται λόγος σαν «το Πανεπιστήμιο» να αντιπροσωπεύει ένα πρόσθετο μέλος της κλάσης όπου ανήκουν ως μέλη οι άλλες μονάδες. Ο επισκέπτης εσφαλμένα κατέτασσε το Πανεπιστήμιο στην ίδια κατηγορία όπου ανήκουν τα άλλα ιδρύματα (Ryle, 1949, p.4).

Το επόμενο διευκρινιστικό παράδειγμα προέρχεται από το χώρο του αθλητισμού. Ένας ξένος παρακολουθεί για πρώτη φορά στη ζωή του αγώνα μπάσκετ³ και μαθαίνει ποιος είναι ο ρόλος των center, του screen, του rinot, των αμυντικών, του διαιτητή, κλπ. Αφού ακούσει τις σχετικές εξηγήσεις, ο επισκέπτης λέει «Όλα καλά, κανείς όμως δεν υπάρχει στο γήπεδο που να ασχολείται με αυτό το σπουδαίο στοιχείο, το ομαδικό πνεύμα. Βλέπω ποιος κάνει τις πάσες, ποιος κάνει το lay up, ποιος μαρκάρει· όμως δεν βλέπω ποιος κάνει το ομαδικό πνεύμα». Ξανά ο επισκέπτης ψάχνει λάθος τύπο πράγματος. «Το ομαδικό πνεύμα δεν είναι μια επιπλέον ενέργεια στο μπάσκετ συμπληρωματική όλων των άλλων ειδικών ενεργειών, αλλά η προθυμία με την οποία επιτελείται κάθε μια από τις ειδικές ενέργειες». Ενώ, είναι σαφές ότι η επίδειξη ομαδικού πνεύματος δεν είναι το ίδιο πράγμα με το μπάσιμο ή την πάσα, την ίδια στιγμή δεν αποτελεί και κάτι διαφορετικό δίπλα σε αυτά με τρόπο που θα μας επέτρεπε να πούμε ότι ο αμυντικός πρώτα μαρκάρει και μετά επιδεικνύει ομαδικό πνεύμα ή ότι ο κυνηγός σε δεδομένη στιγμή είτε σουτάρει για τρίποντο είτε επιδεικνύει ομαδικό πνεύμα. Τα παραπάνω διευκρινιστικά παραδείγματα σφαλμάτων κατηγορίας έχουν ένα αξιοσημείωτο κοινό στοιχείο, που έχει να κάνει με τη λανθασμένη χρήση εννοιών και την αδυναμία κατανόησης του ρόλου τους στη γλώσσα (Ryle, 1949, p.5).

³ Το παράδειγμα του μπάσκετ χρησιμοποιήθηκε από τον Νίκο Μαραγκό κατά τη μετάφραση του κειμένου του Gilbert Ryle, εν αντιθέσει του παραδείγματος του κρίκετ που βρίσκεται στο πρωτότυπο.

Στην πηγή της θεωρίας της διπλής ζωής - μιας νοητικής, μιας φυσικής - βρίσκεται για τον Ryle μια οικογένεια κατηγοριακών σφαλμάτων. Με το επιχείρημα αυτό σχηματοποιεί την εικόνα ότι οι άνθρωποι είναι σαν φαντάσματα που μυστηριωδώς κατοικούν σε μηχανές. Ο φιλόσοφος δέχεται, ωστόσο, ότι η ανθρώπινη νόηση, οι αισθήσεις και οι πράξεις που γίνονται με σκοπιμότητα δεν δύναται να περιγραφούν αποκλειστικά μέσω των επιστημών της χημείας, της φυσικής ή της φυσιολογίας. Προκύπτει έτσι η ανάγκη να αναζητηθούν αλλού οι απαραίτητες περιγραφές.

Ο Ryle διερευνά τα αίτια αυτού του σφάλματος, σκιαγραφώντας τις συνθήκες που διαμόρφωσαν τη σκέψη του Καρτέσιου και τον ώθησαν να υιοθετήσει τη συγκεκριμένη θεωρία. Όταν ο Γαλιλαίος διατύπωσε τη μηχανική θεωρία του για την περιγραφή του χώρου, ο Καρτέσιος βρέθηκε σε μια αμήχανη θέση· ενώ όφειλε σαν επιστήμονας να δεχθεί την παραπάνω θεωρία, παράλληλα σαν θρησκευόμενος και ηθικός άνθρωπος δεν μπορούσε να δεχτεί ότι η ανθρώπινη φύση διαφέρει από το μηχανισμό των ρολογιών μόνο ως προς το βαθμό πολυπλοκότητας. Το διανοητικό δεν μπορούσε να είναι απλώς μια ακόμα πτυχή του μηχανικού.

Η λύση στην οποία διέφυγαν ο Καρτέσιος και άλλοι φιλόσοφοι περνούσε μέσα από αυτή τη διχοτόμηση. Ο Ryle εξηγεί ότι «αφού οι λέξεις οι σχετικές με τη νοητική διαγωγή δεν πρέπει να εκλαμβάνονται ως σημαίνουσες την πραγμάτωση μηχανικών διεργασιών, πρέπει να εκλαμβάνονται ως σημαίνουσες την πραγμάτωση μη μηχανικών διεργασιών. Αφού οι μηχανικοί νόμοι εξηγούν τις χωρικές κινήσεις ως αποτελέσματα άλλων χωρικών κινήσεων, άλλοι νόμοι, διαφορετικής φύσεως, πρέπει να εξηγούν τις μη χωρικές διεργασίες του νου ως αποτελέσματα άλλων μη χωρικών νοητικών διεργασιών» (Ryle, 1949, p.6).

Οι διαφορές ανάμεσα στο φυσικό και στο νοητικό παρουσιάζονταν έτσι ως διαφορές μέσα στο κοινό πλαίσιο των κατηγοριών «πράγμα», «ουσία», «ιδιότητα», «αίτιο», και «αποτελεσμα». Ο νους είναι πράγμα, αλλά διαφορετικού είδους πράγμα από το σώμα· οι νοητικές διεργασίες είναι αίτια και αποτελέσματα, αλλά διαφορετικού είδους αίτια και αποτελέσματα από τις σωματικές κινήσεις. Όπως ο ξένος επισκέπτης περίμενε να είναι το Πανεπιστήμιο ένα επιπλέον κτίριο, όμοιο εν μέρει με κολέγιο αλλά και σημαντικά διαφορετικό, έτσι και οι πολέμιοι του μηχανισμού παρουσίαζαν τον νου σαν ένα επιπλέον κέντρο αιτιακών διεργασιών, εν μέρει όμοιο με μηχανή αλλά και σημαντικά διαφορετικό. Η θεωρία τους ήταν μια παραμηχανική υπόθεση (Ryle, 1949, p.6).

Η παραπάνω παραδοχή ήρθε ως απάντηση στη θεωρητική δυσκολία που προέκυπτε από το δόγμα της αλληλεπίδρασης νου-σώματος. Πώς μπορεί μια νοητική διεργασία, να είναι αίτιο χωρικών κινήσεων όπως των κινήσεων του χεριού; Πώς μπορεί μια φυσική αλλαγή στο οπτικό νεύρο να προκαλέσει τη νοητική αντίληψη μιας φωτεινής αναλαμπής; Ο Καρτέσιος προσπαθώντας να δώσει μια ικανοποιητική λύση στα παραπάνω, χρησιμοποίησε τη γραμματική της γαλιλαϊκής μηχανικής, απλά αυτή τη φορά ανεστραμμένη. Οι περιγραφές των νοητικών συμβάντων συνιστούν πλέον την άρνηση των ειδικών περιγραφών που αφορούν τα σώματα· το τι γίνεται στον νου δεν βρίσκεται στο χώρο, δεν διακρίνεται από την ιδιότητα της έκτασης, δεν είναι τροποποίηση της ύλης, δεν είναι προσιτό στη δημόσια παρατήρηση.

«Θεωρούμενος κατ'αυτόν τον τρόπο», συμπληρώνει ο Ryle, «ο νους δεν είναι απλώς ένα φάντασμα δεμένο σε μια μηχανή, είναι ο ίδιος μια φασματική μηχανή και τίποτα άλλο». Μια διαφορετική, εσωτερική μηχανή κυβερνά δίνοντας εντολές, τη μηχανή του σώματος. Για αυτή την αόρατη, αθόρυβη μηχανή και τον τρόπο λειτουργίας της λίγα πράγματα γνωρίζουμε καθώς «δεν μπορεί να αναλυθεί σε μέρη και την διέπουν νόμοι άγνωστοι στους συνηθισμένους μηχανικούς» (Ryle, 1949, p.7).

Μια δεύτερη σημαντική δυσκολία έγκειται στο ζήτημα της ελευθερίας της βούλησης. Αφού, κατά το δόγμα, ο νους ανήκει στην ίδια κατηγορία με το σώμα και αφού το σώμα το κυβερνούν μηχανικοί νόμοι, ο νοητικός, όπως και ο φυσικός κόσμος, πρέπει να αποτελεί αιτιοκρατικό σύστημα. Τόσο η πορεία των σωματικών κινήσεων, χειρονομιών, πράξεων όσο και των νοητικών λειτουργιών που αφορούν τη βούληση, την επιθυμία, την πίστη, είναι καθορισμένες εκ των προτέρων. Έτσι δεδομένου ότι οι νόμοι που διέπουν τις νοητικές διεργασίες είναι εξίσου δύσκαμπτοι με τους φυσικούς νόμους, έννοιες όπως ευθύνη, επιλογή και ελευθερία καθίστανται χωρίς νόημα. «Το πρόβλημα της ελευθερίας της βούλησης ήταν το εξής: πώς να συμβιβαστεί η υπόθεση ότι ο νους πρέπει να περιγράφεται με όρους αντλημένους από τις κατηγορίες της μηχανικής αφενός, με τη γνώση ότι η ανώτερης τάξης ανθρώπινη διαγωγή δεν είναι της ίδιας τάξης με τη συμπεριφορά των μηχανών αφετέρου» σημειώνει ο Ryle (Ryle, 1949, p.7).

Το σφάλμα κατηγορίας που γεννά το φάντασμα της μηχανής βασίζεται στην εξής γλωσσολογική απάτη. Για τον Καρτέσιο, υπάρχει το σώμα και ο νους, φυσικές και νοητικές διεργασίες, μηχανικά αίτια σωματικών κινήσεων και νοητικά αίτια σωματικών κινήσεων. Για το Ryle τέτοιου είδους συμπλεκτικές προτάσεις είναι παράλογες, αν και όχι παράλογες καθ'αυτές. «Δεν αρνούμαι, π.χ., ότι συντελούνται νοητικές διεργασίες. Η διαίρεση πολυψήφιων αριθμών είναι νοητική διεργασία όπως και η επινόηση ενός αστείου. Λέω όμως ότι η φράση 'συντελούνται νοητικές διεργασίες' δεν σημαίνει το ίδιο πράγμα με τη φράση 'συντελούνται φυσικές διεργασίες', και ως εκ τούτου δεν έχει νόημα η σύζευξη ή η διάζευξη τους» (Ryle, 1949, p.8).

Η σκέψη του Ryle στοχεύει στη διάλυση της «καθαγιασμένης αντίθεσης» ανάμεσα στο νου και το σώμα, αλλά όχι με τρόπο που να αναγάγει το ένα στο άλλο. Αυτή η φαινομενική αντίθεση διαλύεται αν θεωρήσουμε τις φράσεις «περπατάω με βροχή» και «περπατάω με τους φίλους μου»: η σύζευξη ή διάζευξη των δυο φράσεων μοιάζει παράλογη. Οι φαινομενικές αντιθέσεις βασίζονται στη λανθασμένη πεποίθηση ότι «νους» και «ύλη» αποτελούν όρους του ίδιου λογικού τύπου (Ryle, 1949, p.7).

Για τον Ryle, ο ιδεαλισμός και ο υλισμός είναι απαντήσεις σε λάθος ερώτηση. Τόσο η αναγωγή υπέρ του υλισμού όσο και η αναγωγή υπέρ του ιδεαλισμού, μας υποχρεώνει να δεχτούμε τη διάζευξη «υπάρχει νους ή υπάρχει σώμα (αλλά όχι και τα δύο)». Στο παράδειγμά θα λέγαμε, «περπατάω με τους φίλους μου ή περπατάω με βροχή (αλλά όχι και τα δύο)» (Ryle, 1949, p.7).

«Είναι απολύτως ορθό να λέμε, με τον ένα λογικό τόνο φωνής, ότι υπάρχει νους, και, με έναν άλλο λογικό τόνο φωνής, να λέμε ότι υπάρχει σώμα» προσθέτει ο Ryle. Δεν πρόκειται όμως για διαφορετικά είδη ύπαρξης, καθώς οι όροι 'νους' και 'σώμα'

υποδεικνύουν απλά δύο διαφορετικές σημασίες του όρου ‘υπάρχει’. Αντίστοιχα ο όρος «ανεβαίνει» αποκτά διαφορετικές σημασίες στις φράσεις «η παλίρροια ανεβαίνει», «η διάθεση ανεβαίνει» και «το προσδόκιμο όριο επιβίωσης ανεβαίνει». Για τον Ryle -όσο και για μας- «θα ήταν αποτυχημένο το αστείο ότι ανεβαίνουν τρία πράγματα, η παλίρροια, η διάθεση και το προσδόκιμο όριο επιβίωσης. Θα ήταν εξίσου αμφίβολο αστείο αν κάποιος έλεγε ότι υπάρχουν πρώτοι αριθμοί και Τετάρτες και κοινές γνώμες και πολεμικά ναυτικά· ή ότι υπάρχει και ο νους και το σώμα» (Ryle, 1956, p.9).

Κεφάλαιο 3 : Η φιλοσοφική ματιά του Roger Penrose στα ερωτήματα της συνείδησης

3.1 Εισαγωγή στη σκέψη του R.Penrose

Στο βιβλίο του «Σκιές του Νου» (Shadows of the Mind), ο Roger Penrose παρουσιάζει ένα επιχείρημα που αμφισβητεί τα επικρατούντα υπολογιστικά μοντέλα της συνείδησης. Ο Penrose, διάσημος μαθηματικός και φυσικός, φέρνει μια μοναδική προοπτική στη μελέτη της συνείδησης, αξιοποιώντας την εμπειρία του από την κβαντομηχανική και τη μαθηματική συλλογιστική. Η κεντρική θέση του Penrose είναι ότι η συνείδηση δεν μπορεί να εξηγηθεί πλήρως ή να αναχθεί σε υπολογιστικές διαδικασίες. Υποστηρίζει ότι υπάρχουν πτυχές της ανθρώπινης νόησης, ιδίως στη δημιουργική και διαισθητική σκέψη, που υπερβαίνουν τις δυνατότητες οποιουδήποτε αλγοριθμικού ή υπολογιστικού συστήματος. Ο νους για τον Penrose διαθέτει ένα θεμελιωδώς μη υπολογιστικό στοιχείο που διαδραματίζει κρίσιμο ρόλο στην παραγωγή της συνειδητής εμπειρίας. Αντλώντας από τους περιορισμούς των τυπικών συστημάτων, όπως το θεώρημα της μη πληρότητας που ανέπτυξε ο Kurt Gödel - σύμφωνα με το οποίο υπάρχουν μαθηματικές αλήθειες που δεν μπορούν να αποδειχθούν στο πλαίσιο οποιουδήποτε τυπικού συστήματος - ο συγγραφέας υποστηρίζει ότι ο ανθρώπινος νους διαθέτει παρόμοια ικανότητα να συλλαμβάνει και να κατανοεί αλήθειες που ξεπερνούν την εμβέλεια των υπολογιστικών αλγορίθμων (Penrose R. , 1996).

Στο πρώτο μέρος του βιβλίου ο συγγραφέας αναπτύσσει τη θέση ότι το φαινόμενο της συνείδησης μπορεί να εμφανιστεί μόνο μέσα από κάποιες μη υπολογιστικές φυσικές διαδικασίες που λαμβάνουν χώρα εντός του εγκεφάλου. Για να το πετύχει αυτό, εγείρει ενστάσεις απέναντι σε διαφορετικά επιχειρήματα αναγωγιστικών προσεγγίσεων φέρνοντας παραδείγματα τόσο από τη λογική, τα μαθηματικά αλλά και τις σύγχρονες θεωρίες για την τεχνητή νοημοσύνη.

Εκκινώντας από την απόρριψη των αυστηρά υπολογιστικών προσεγγίσεων, στη συνέχεια του βιβλίου προχωρά προς την αναζήτηση μιας νέας φυσικής για την κατανόηση της νόησης. Σε αυτήν την κατεύθυνση, παρατηρούμε τον Penrose να εισάγει την έννοια των *μικροσωληνίσκων* εντός των νευρώνων ως πιθανή πηγή μη υπολογιστικών διαδικασιών. Αυτές οι δομές παρουσιάζουν κβαντομηχανική συμπεριφορά και μπορεί να παίζουν ρόλο στη δημιουργία συνείδησης μέσω της θεωρίας του της *Ενορχηστρωμένης Αντικειμενικής Αναγωγής* (Orchestrated Objective Reduction ή Orch-OR). Σύμφωνα με αυτή τη θεωρία, η συνείδηση προκύπτει από τα κβαντικά φαινόμενα εντός των μικροσωληνίσκων, τα οποία είναι υπεύθυνα για την εμφάνιση της υποκειμενικής εμπειρίας (Penrose & Hameroff, 1996).

Το επιχείρημα του Penrose αμφισβητεί την άποψη ότι η συνείδηση είναι αποκλειστικά προϊόν υπολογιστικών διεργασιών στον εγκέφαλο. Τονίζοντας τους περιορισμούς των υπολογιστικών μοντέλων και δείχνοντας προς μια κατεύθυνση κατανόησης των συνειδητών λειτουργιών μέσω κβαντομηχανικών εργαλείων και εννοιών,

ο Penrose προσφέρει ένα εναλλακτικό πλαίσιο για την κατανόηση των υποκείμενων μηχανισμών της συνείδησης.

3.2 Μη υπολογιστική προσέγγιση της συνείδησης

Στα εισαγωγικά κεφάλαια του βιβλίου, ο Penrose μας φέρνει σε επαφή με την κατηγοριοποίησή του για τις τέσσερις διαφορετικές θέσεις σε σχέση με το ζήτημα της υπολογισιμότητας της συνείδησης, οι οποίες παρουσιάζονται συνοπτικά παρακάτω:

- A. *Ισχυρή τεχνητή νοημοσύνη (strong AI - functionalism)*: Η άποψη αυτή υποστηρίζει ότι η συνείδηση είναι πλήρως υπολογίσιμη και μπορεί να προσομοιωθεί από υπολογιστικές διαδικασίες⁴. Ισχυρίζεται ότι ο νους είναι ουσιαστικά ένα σύστημα επεξεργασίας πληροφοριών και οι συνειδητές εμπειρίες προκύπτουν από την εκτέλεση αλγορίθμων στο φυσικό υπόστρωμα του εγκεφάλου. Βασική συνέπεια της θέσης είναι η υπόθεση ότι ένα επαρκώς προηγμένο πρόγραμμα υπολογιστή θα μπορούσε να εκδηλώσει συνειδητή συμπεριφορά.
- B. *Ασθενής Τεχνητή Νοημοσύνη (weak AI - connectionism)*: Αυτή η προοπτική υιοθετεί μια πιο μετριοπαθή θέση σε σχέση με την Α. Η ασθενής ΤΝ, γνωστή και ως θεωρία νευρωνικών δικτύων, αναγνωρίζει το ρόλο του υπολογισμού στη νόηση, αλλά πιστεύει ότι από μόνος του δεν επαρκεί για να εξηγήσει πλήρως τις συνειδητές εμπειρίες. Τονίζει τη σημασία της παράλληλης κατανεμημένης επεξεργασίας και των νευρωνικών δικτύων στη γνωστική λειτουργία, υποδηλώνοντας ότι η συνείδηση προκύπτει από τις πολύπλοκες αλληλεπιδράσεις των διασυνδεδεμένων νευρωνικών μονάδων. Η συνειδητότητα παραμένει χαρακτηριστικό της φυσικής λειτουργίας του εγκεφάλου· «ενώ όμως οποιαδήποτε φυσική δραστηριότητα μπορεί να προσομοιωθεί σε υπολογιστή, καμία προσομοίωση δεν μπορεί να δημιουργήσει συνειδητότητα» (Penrose R., 1996, p.12).
- C. *Φυσική μη υπολογιστική διαδικασία*: Αυτή η κατηγορία υποδηλώνει ότι η συνείδηση περιλαμβάνει μη υπολογίσιμες διεργασίες που υπερβαίνουν αυτό που μπορεί να

⁴ Με τον όρο υπολογιστική διαδικασία (ή αλγόριθμος) -όροι που χρησιμοποιούνται εναλλάξιμα από τον Penrose- νοείται η δραστηριότητα κάποιας μηχανής Turing, η λειτουργία, δηλαδή, ενός υπολογιστή βάσει κάποιου ειδικού προγράμματος. Οι υπολογιστικές διαδικασίες δεν περιορίζονται στις συνηθισμένες πράξεις της αριθμητικής όπως πρόσθεση και πολλαπλασιασμός, αλλά περιλαμβάνουν και καλά ορισμένες λογικές πράξεις. Πιο συγκεκριμένα, ένας αλγόριθμος αποτελεί μια διαδικασία που μπορεί να εκτελεστεί από μια μηχανή Turing, την οποία μπορούμε να φανταστούμε ως ένα μαθηματικά εξειδικευμένο υπολογιστή. Μια τέτοια μηχανή εκτελεί τις δραστηριότητές της με μια βήμα προς βήμα διαδικασία, στην οποία κάθε βήμα καθορίζεται απόλυτα από τη φύση κάποιων συμβόλων πάνω σε μια ταινία που η μηχανή εξετάζει κάθε χρονική στιγμή και από την «εσωτερική κατάσταση» της μηχανής. Οι διάφορες επιτρεπτές καταστάσεις είναι πεπερασμένου πλήθους, όπως και ο συνολικός αριθμός των συμβόλων πάνω στην ταινία. Η μηχανή ξεκινά από μια συγκεκριμένη κατάσταση, π.χ. από την κατάσταση «0», και οι εντολές δίνονται στην ταινία με τη μορφή δυαδικών αριθμών. Διαβάζει τις εντολές, μετατοπίζοντας την ταινία σύμφωνα με τις βηματικές της εσωτερικές διαδικασίες και με τρόπο που καθορίζεται σε κάθε στάδιο από την εσωτερική της κατάσταση και το συγκεκριμένο στοιχείο που εξετάζει. Η μηχανή μπορεί να σβήσει υπάρχοντα ή να δημιουργήσει νέα σύμβολα. Όταν συναντήσει την εντολή STOP, εμφανίζει στην ταινία το αποτέλεσμα του υπολογισμού και σταματά μέχρι την πραγματοποίηση του επόμενου.

επιτευχθεί μέσω αλγοριθμικού υπολογισμού. Η συνειδητότητα προκαλείται από κατάλληλη φυσική δραστηριότητα του εγκεφάλου, αλλά αυτή η δραστηριότητα δεν είναι προσομοιώσιμη σε υπολογιστή.

- D. Η εμπειρία της συνείδησης αποτελεί υπερβατικό/μεταφυσικό φαινόμενο καθώς αυτή δεν μπορεί να περιγραφεί με φυσικούς, υπολογιστικούς ή άλλους επιστημονικούς όρους.

Στο επιχείρημά του, ο Penrose απορρίπτει τη θεωρία των υπολογιστικών μοντέλων ως επαρκή εξήγηση για τη συνείδηση. Αμφισβητεί την επικρατούσα αντίληψη ότι οι λειτουργίες της νόησης μπορούν να εξηγηθούν πλήρως από αλγοριθμικές διαδικασίες και υπολογιστικές προσομοιώσεις. Ο συγγραφέας υποστηρίζει ότι τα υπολογιστικά μοντέλα, συμπεριλαμβανομένων των συστημάτων τεχνητής νοημοσύνης (AI), έχουν εγγενείς περιορισμούς στην αποτύπωση του πλούτου και της πολυπλοκότητας της ανθρώπινης συνείδησης. Υποστηρίζει ότι η συνειδητή ανθρώπινη σκέψη περιλαμβάνει περισσότερα από την εκτέλεση αλγορίθμων ή τον χειρισμό συμβόλων. Σύμφωνα με τον Penrose, υπάρχουν πτυχές της ανθρώπινης νόησης, όπως η διαίσθηση και η δημιουργικότητα, που δεν μπορούν να αναχθούν σε υπολογιστικές διαδικασίες. Μία από τις βασικές προκλήσεις που θέτει ο Penrose για τα υπολογιστικά μοντέλα είναι το πρόβλημα της κατανόησης και της παραγωγής μαθηματικής σκέψης. Υποστηρίζει ότι η μαθηματική κατανόηση περιλαμβάνει μια μη αλγοριθμική συνιστώσα, όπως αποδεικνύεται από το θεώρημα της μη πληρότητας του Gödel, καθώς ο ανθρώπινος νους έχει την ικανότητα να αντιλαμβάνεται μαθηματικές αλήθειες που ξεπερνούν τις δυνατότητες οποιουδήποτε υπολογιστικού συστήματος (Penrose R. , 1996).

Αμφισβητώντας την επάρκεια των υπολογιστικών μοντέλων, ο Penrose ανοίγει έτσι τη δυνατότητα εναλλακτικών εξηγήσεων για τη συνείδηση. Τα επιχειρήματά του προσκαλούν σε περαιτέρω διερεύνηση των μη υπολογιστικών στοιχείων στη νόηση και του δυναμικού ρόλου της κβαντομηχανικής στην κατανόηση της συνειδητής εμπειρίας. Η απόρριψη των υπολογιστικών μοντέλων από τον Penrose ακολουθείται από μια βαθύτερη εξέταση της Orch-OR θεωρίας του, η οποία προτείνει έναν μη υπολογιστικό μηχανισμό για τη συνείδηση. Θέτει έτσι τις βάσεις για την αξιολόγηση των συνεπειών της μη υπολογιστικής προσέγγισης του Penrose και της συμβατότητάς της με τα εμπειρικά δεδομένα και τις εναλλακτικές προοπτικές στη μελέτη της συνείδησης.

3.2.1 Σχετικά με το δυϊσμό σώματος-νου

Γιατί δεχόμαστε πλέον, καθολικά, το γεγονός ότι τα σώματά μας υπακούνε στους περιορισμούς των φυσικών νόμων, αλλά αδυνατούμε να αποδεχτούμε το ίδιο και για το πνεύμα μας; Ο ισχυρισμός αυτός φέρνει σε αμηχανία πολλούς ανθρώπους: η διαπίστωση, δηλαδή, ότι η νόησή μας μπορεί επίσης να είναι αναγκασμένη να δρα σύμφωνα με τους ίδιους μαθηματικούς νόμους. Πιο ανησυχητική, βρίσκει ο Penrose, τη χάραξη μιας διαχωριστικής γραμμής μεταξύ σώματος και νου, με το πρώτο να είναι «δεμένο» στους

φυσικούς νόμους και το δεύτερο να λειτουργεί εντός ενός μυστικιστικού πλαισίου αόριστης ελευθερίας. «Γιατί το πνεύμα μας», δηλώνει ο συγγραφέας, «επηρεάζει σίγουρα τον τρόπο με τον οποίο δρα το σώμα μας, και πρέπει επίσης να επηρεάζεται από τη φυσική κατάσταση του ίδιου του σώματος. Αν το πνεύμα μπορούσε να επηρεάζει το σώμα, επιτρέποντάς του μεγαλύτερους βαθμούς ελευθερίας από αυτούς που θέτουν οι φυσικοί νόμοι, αυτό θα υπέσκαπτε την αυστηρότητα και την ακρίβεια των επιστημονικών νόμων». Από την άλλη, η θέση που θέλει το πνεύμα να είναι απλώς ένα επιφανόμενο -κάποιο ειδικό, παθητικό στοιχείο της φυσικής λειτουργίας του εγκεφάλου- το οποίο είναι απλά ένα υποπροϊόν του σώματος χωρίς επίδραση πάνω σε αυτό, υποβιβάζει το πνεύμα δίνοντας του έναν αδύναμο και απογοητευτικό ρόλο. Ακόμα και αν οι φυσικοί νόμοι επιτρέπουν κάποια ελευθερία στο σώμα, στα πλαίσια της οποίας το πνεύμα μπορεί να επηρεάσει τη συμπεριφορά του σώματος, η ίδια η φύση αυτής της ελευθερίας πρέπει να αποτελεί ένα σημαντικό συστατικό αυτών των νόμων. Μπορούμε να ερμηνεύσουμε αυτή την χειρονομία του Penrose μέσω της προοπτική της *ριζικής εμμένειας* σε σχέση με τη φύση της συνείδησης (βλ. Μπαλτάς). «Οτιδήποτε και αν είναι αυτό που καθορίζει ή περιγράφει το πνεύμα, θα πρέπει πράγματι να αποτελεί ένα αναπόσπαστο κομμάτι του ίδιου μεγαλειώδους σχήματος που κυβερνά, επίσης, όλα τα *υλικά* χαρακτηριστικά του σύμπαντός μας» (Penrose R., 1996, p.213).

Ο Penrose δεν αποφεύγει να σχολιάσει την κριτική που του γίνεται σε σχέση με τη διάπραξη κατηγορικού σφάλματος στην ανάπτυξη της φιλοσοφικής του θεώρησης. Η κριτική βασίζεται στην ιδέα ότι ο Penrose αναφέρεται στο «πνεύμα» σαν αυτό να είναι ένα άλλο είδος ουσίας, διαφορετικό από την ύλη και οριζόμενο από άλλες αρχές, υποπίπτοντας, έτσι, σε σφάλμα κατηγορίας, με αποτέλεσμα να δημιουργείται σύγχυση μεταξύ διαφορετικών ειδών εννοιών. Φέρνοντας ως παράδειγμα την εξίσωση του Einstein για την ισοδυναμία μάζας ενέργειας, σημειώνει το φαινομενικό σφάλμα κατηγορίας που θα μπορούσε να προκύψει από την εξίσωση μιας υλικής ουσίας με μια πιο «νεφελώδη αφηρημένη ποσότητα», τη δυνατότητα παραγωγής έργου. Η ευρεία της πειραματική εφαρμογή και επαλήθευση παρ'όλ'αυτά, την καθιστά ένα από τα σημαντικότερα εργαλεία της σύγχρονης Φυσικής. Αντίστοιχα, θεωρεί απόλυτα «νόμιμη» μια συζήτηση που να συσχετίζει το πνεύμα με άλλες φυσικές έννοιες. Αναφερόμενος στη συνείδηση, συνεχίζει, «φαίνεται ότι είναι κάτι που «υπάρχει» σε συνάρτηση με κάποια ειδικά φυσικά αντικείμενα, με τους ζώντες ενεργούς ανθρώπινους εγκεφάλους, τουλάχιστον» προσδοκώντας σε μια πιο ολοκληρωμένη φυσική περιγραφή του φαινομένου στο μέλλον (Penrose R., 1996, p.214).

3.3 Το θεώρημα του Gödel

Στα μαθηματικά μπορούμε να εξετάσουμε τις διαδικασίες της σκέψης μας στην καθαρότερη μορφή τους. Η διάθεση του Penrose να εστιάσει στον τομέα των μαθηματικών στην αναζήτησή του για απαντήσεις στα ερωτήματα σχετικά με την ανθρώπινη νόηση, προκύπτει από την υπόθεση ότι «αν σκέψη είναι απλώς η πραγματοποίηση ενός υπολογισμού οποιουδήποτε είδους, φαίνεται ότι θα πρέπει να είμαστε σε θέση να το διαπιστώσουμε περισσότερο ξεκάθαρα, στη μαθηματική σκέψη» (Penrose R., 1996, p.64). Ωστόσο ο Penrose ισχυρίζεται ότι μόνο εντός του πλαισίου των μαθηματικών εντοπίζονται

οι πιο ξεκάθαρες ενδείξεις ότι οι διαδικασίες της ενσυνείδητης σκέψης υπερβαίνουν κάθε υπολογιστική δυνατότητα.

Το 1930 στο Königsberg, ένας ιδιοφυής νεαρός μαθηματικός ονόματι Kurt Gödel ταρακούνησε τα θεμέλια της μαθηματικής κοινότητας, καταρρίπτοντας τις φιλοδοξίες των οπαδών του φορμαλισμού. Αυτό που έμεινε στην ιστορία σαν το θεώρημα μη πληρότητας του Gödel έγινε αποδεκτό ως μια θεμελιώδης συνεισφορά όχι μόνο στη μαθηματική σκέψη αλλά και στη φιλοσοφία της νόησης. Ανάμεσα στις έννοιες που καθιέρωσε ο Gödel ήταν ότι κανένα τυπικό σύστημα αυστηρών μαθηματικών κανόνων δεν αρκεί, ούτε καν θεωρητικά, για να αποδείξει όλες τις αληθείς προτάσεις της αριθμητικής. Ακόμα περισσότερο, το θεώρημα έδειξε ότι η ανθρώπινη κατανόηση και διαίσθηση δεν μπορούν να αναχθούν σε κανένα σύνολο υπολογιστικών κανόνων. Με άλλα λόγια, κανένα τέτοιο σύστημα δεν μπορεί να είναι αρκετό να αποδείξει ακόμα και εκείνες τις προτάσεις της αριθμητικής των οποίων η αλήθεια είναι προσεγγίσιμη από την ανθρώπινη διαίσθηση, ενώ η ίδια η ανθρώπινη διαίσθηση δεν μπορεί να αναχθεί σε κανένα σύνολο κανόνων. Ο Penrose τονίζει τον εξέχοντα ρόλο που διαδραματίζουν οι συνέπειες του θεωρήματος στις ιδέες του, καθώς αυτό αποτελεί ακρογωνιαίο λίθο της συλλογιστικής που χρησιμοποιεί (Penrose R., 1996, p.213).

3.4 Η χρονική ροή στον Penrose

Κατά γενική ομολογία το *πέραςμα του χρόνου* αποτελεί ένα από τα πιο άμεσα χαρακτηριστικά της ενσυνείδητης αντίληψης. Αξιοσημείωτο είναι το γεγονός ότι παρά το πλούτο που μας έχουν προσφέρει οι επιστημονικές θεωρίες σε σχέση με την κατανόηση της συμπεριφοράς του φυσικού κόσμου, δεν έχουν καταφέρει να δώσουν πειστικές απαντήσεις σε σχέση με το παραπάνω ζήτημα. Αντίθετα, όσα έχουν ειπωθεί φαίνονται να έρχονται σε απόλυτη αντίθεση με την καθημερινή μας αντίληψη για το χρόνο.

Για τη γενική σχετικότητα, ο χρόνος είναι απλά μια συγκεκριμένη επιλογή συντεταγμένης για την περιγραφή της θέσης ενός χωροχρονικού γεγονότος. Δεν υπάρχει κάτι στις χωροχρονικές περιγραφές που να διαφοροποιεί το χρόνο ως κάτι που ρέει. Ακόμα όμως, κι αν σε δισδιάστατους χωροχρόνους δεν υπάρχει κάτι που μας επιτρέπει να διακρίνουμε τις δύο διαφορετικές διαστάσεις, δεν μπορεί να γίνει λόγος για κάποιον χώρο που «ρέει». Η χρονική εξέλιξη λαμβάνεται συχνά υπόψη σε φυσικά προβλήματα, όπου μπορεί να ενδιαφέρει ο υπολογισμός του μέλλοντος με βάση την παρούσα κατάσταση του προβλήματος. Εντούτοις, μια τέτοια διαδικασία δε φαίνεται απαραίτητη από φυσική άποψη. Οι υπολογισμοί πραγματοποιούνται με αυτόν τον τρόπο επειδή μας ενδιαφέρει να δημιουργήσουμε ένα μαθηματικό μοντέλο των εμπειριών μας από τον κόσμο με βάση το ρέοντα χρόνο που φαίνεται να αντιλαμβανόμαστε, ώστε να μπορούμε να κάνουμε προβλέψεις για το μέλλον. Οι φυσικοί νόμοι -σε αντίθεση με τις εμπειρίες στις οποίες βασίζουμε τα μοντέλα μας- δεν περιέχουν καμιά τέτοια προκατάληψη.

Μόνο το φαινόμενο της συνείδησης απαιτεί από μας να σκεφτόμαστε με βάση ένα ρέοντα χρόνο. Σύμφωνα με τη γενική σχετικότητα, υπάρχει ένας στατικός τετραδιάστατος χωροχρόνος, ο οποίος «είναι απλά εκεί» ενώ ο χρόνος δε ρέει περισσότερο από ότι ο χώρος.

Είναι προφανές λοιπόν, ότι μόνο η συνείδηση μας χρειάζεται κάποια έννοια περάσματος του χρόνου.

Ο Penrose αναρωτιέται για τη δυνατότητα να ελέγξουμε πειραματικά πότε μια υποκειμενική εμπειρία πράγματι συνέβη, σε σχέση με τη φυσική παράμετρο του χρόνου t που χρησιμοποιούν οι φυσικοί στις περιγραφές τους. Ποια η αντικειμενική σημασία της πρότασης «ένα συνειδητό γεγονός συμβαίνει σε μια συγκεκριμένη χρονική στιγμή;» (Penrose R., 1996, p.384). Σε αναζήτηση απαντήσεων ο Penrose ανατρέχει σε σχετικά πειράματα με φαινομενικά αρκετά παράδοξα αποτελέσματα.

Στην προσπάθεια τους να χρονομετρήσουν οποιαδήποτε εγκεφαλική δραστηριότητα που να συνδέεται με κάποια πράξη ελεύθερης βούλησης, μια ομάδα επιστημόνων στα μέσα της δεκαετίας του 1970 χρησιμοποίησε ηλεκτροεγκεφαλογράφημα για να καταγράψει ηλεκτρικά σήματα σε διαφορετικά σημεία στα κεφάλια εθελοντών. Από τους εθελοντές ζητήθηκε να λυγίσουν το δείκτη του χεριού τους, ξαφνικά σε στιγμές δικής τους επιλογής, ώστε να χρονομετρηθεί η εγκεφαλική δραστηριότητα που είχε σχέση με την κίνηση του δακτύλου. Το ηλεκτρικό δυναμικό φάνηκε να δημιουργείται σταδιακά από ένα χρονικό διάστημα περίπου ίσο 1-1.5 δευτερόλεπτο, πριν από την πραγματική κάμψη του δακτύλου. Οδηγούμαστε λοιπόν, στο συμπέρασμα ότι μια ενσυνείδητη πράξη της βούλησης χρειάζεται ένα δευτερόλεπτο ή περισσότερο για να πραγματοποιηθεί; Αν ο εθελοντής είχε πλήρη συνείδηση, απόφαση του να λυγίσει το δάκτυλό του θα έπρεπε να ληφθεί μόλις μια στιγμή πριν το λυγίσει, και σίγουρα όχι τόσο πριν όσο το χρονικό διάστημα ενός δευτερολέπτου (Penrose R., 1996, p.386).

Για τον Penrose το παραπάνω πείραμα δείχνει καταρχάς ότι η ενσυνείδητη δράση της «ελεύθερης βούλησης» είναι μια καθαρή ψευδαίσθηση, αφού «υπό κάποια έννοια, έχει ήδη προ-προγραμματιστεί από την προηγούμενη ασυνείδητη δραστηριότητα του εγκεφάλου» (Penrose R., 1996, p.386). Ή πάλι, θα μπορούσε να υπάρχει ένας πιθανός ρόλος της «τελευταίας στιγμής» για τη βούληση, έτσι ώστε να είναι δυνατή η αντιστροφή της απόφασης που λαμβάνεται ασυνείδητα μέσα στο χρονικό διάστημα του περασμένου δευτερολέπτου. Θα μπορούσε ακόμα, ο εθελοντής πράγματι να επιθυμεί συνειδητά το λύγισμα του δακτύλου του ένα δευτερόλεπτο πριν το πραγματοποιήσει, αλλά να αντιλαμβάνεται λανθασμένα ότι η ενσυνείδητη απόφαση έχει ληφθεί πολύ αργότερα, μόλις λίγο πριν το λύγισμα του δακτύλου.

«Είναι δυνατό η φυσική επιλογή» σημειώνει ο Penrose, «να έχει δημιουργήσει τη συνείδηση αποκλειστικά για το ρόλο που παίζει στην προμελετημένη σκέψη, ενώ σε οποιαδήποτε ταχεία δραστηριότητα η συνείδηση είναι απλά ένας επιβάτης» (Penrose R., 1996, p.387). Η μαθηματική κατανόηση για την οποία έγινε λόγος προηγουμένως, αποτελεί τέτοιο είδος ενσυνείδητου συλλογισμού, δηλαδή βραδύ συλλογισμού για τις ανάγκες του οποίου ίσως αναπτύχθηκε η δύναμη της συνείδησης. Ταυτόχρονα, οι ταχύτεροι χρόνοι απόκρισης σε διεγέρσεις αφορούν το ασυνείδητο κομμάτι της εγκεφαλικής δραστηριότητας, συνοδευόμενοι από μια καθυστερημένη ενσυνείδητη αντίληψη που δε διαδραματίζει κάποιο ρόλο στη γενικότερη λειτουργία.

Ο Penrose αναγνωρίζει την ανάγκη κάποιου σχετικά μεγάλου χρονικού διαστήματος για την ανάπτυξη της συνείδησης, αφήνει όμως ανοιχτό το ενδεχόμενο η συνείδηση να παίζει σημαντικό ρόλο σε τέτοιες ταχείες δραστηριότητες όπως ένας καθημερινός διάλογος, ένα παιχνίδι τένις κλπ. Υπάρχει, σύμφωνα με το συγγραφέα, «μια εγγενής ασάφεια σε ό,τι αφορά τη σχέση ανάμεσα στην ενσυνείδητη εμπειρία και το φυσικό χρόνο» (Penrose R., 1996, p.388), ενώ ειδικά στις περιπτώσεις που υπεισέρχονται κβαντικά φαινόμενα, τα φαινομενικά λογικά συμπεράσματα που αφορούν τη χρονική διάταξη των γεγονότων τίθενται υπό αμφισβήτηση. Αντίστροφα, αν κατά την εκδήλωση της συνείδησης, οι κλασσικοί συλλογισμοί μας για τη χρονική διαδοχή, την αιτιότητα και την τοπικότητα οδηγήσουν σε λογικά αδιέξοδα, οι επιστήμονες πρέπει να εστιάσουν σε μια αναθεώρηση του ρόλου της κβαντικής θεωρίας.

3.5 Τα όρια της Τεχνητής Νοημοσύνης

Ευθύς εξαρχής, ο Penrose τάσσεται υπέρ της θέσης C ασκώντας κριτική αμφότερα στις διαφορετικές εκδοχές του AI (strong-weak) και στις πιο μυστικιστικές προσεγγίσεις που αφορούν το φαινόμενο της συνείδησης. Βασική του άποψη αποτελεί η υπόθεση ότι «αυτό που επιτυγχάνουμε με την ενσυνείδητη σκέψη μας είναι πολύ διαφορετικό από αυτό που μπορεί να επιτευχθεί με κάποιες υπολογιστικές διαδικασίες» (Penrose R., 1996, p.393). Χαράσσοντας τα όρια του συγκεκριμένου πεδίου, τοποθετείται ενάντια τόσο στις A και B θεωρήσεις, αμφισβητώντας την δυνατότητα προσομοίωσης της συνειδητής λειτουργίας από έναν υπολογιστή, όσο και στη θεώρηση D, δηλώνοντας απερίφραστα την πεποίθηση του για τη δυνατότητα της επιστημονικής σκέψης να δώσει απαντήσεις σε ερωτήματα για τη φύση των ενσυνείδητων πράξεων. Για να αντιμετωπίσει αυτές τις προκλήσεις προσπαθεί να βρει παραδείγματα νοητικής δραστηριότητας που βρίσκονται πέρα από κάθε είδος υπολογιστικής διαδικασίας, αλλά και να συλλάβει πώς μια τέτοια δραστηριότητα μπορεί να προκύπτει από κατάλληλες φυσικές διαδικασίες.

Στην αναζήτησή του για την εύρεση κάποιας νοητικής δραστηριότητας πέραν κάθε υπολογιστικής διαδικασίας, στρέφεται στην εξέταση της παρούσας κατάστασης της τεχνητής νοημοσύνης. Ενώ κάνει λόγο για την ταχεία ανάπτυξη του τομέα και τις μεγάλες δυνατότητες περαιτέρω προόδου που υπάρχουν, στέκεται κριτικά σε σχέση με κάποιους θεμελιώδεις περιορισμούς αρχής στους οποίους υπόκειται η TN.

Εντοπίζει τα όρια της TN όχι τόσο σε περιοχές που η ανθρώπινη διάνοια μπορεί να είναι εντυπωσιακή, π.χ. σε περιπτώσεις όπου η εξειδίκευση κάποιων ανθρώπων προκαλεί δέος λόγω της ικανότητάς τους να κρίνουν βάση εξαιρετικά περίπλοκων υπολογιστικών διαδικασιών, αλλά περισσότερο σε κοινές, καθημερινές δραστηριότητες στις όποιες επιδιόμαστε κατά το μεγαλύτερο μέρος της ζωής μας. Από την καλλιτεχνική έκφραση και δημιουργικότητα μέχρι τις κοινωνικές αλληλεπιδράσεις και την ανάπτυξη κωδίκων συναισθηματικής αναγνώρισης, υπάρχουν πτυχές της ανθρώπινης εμπειρίας και ικανότητας που αποτελούν ακόμα πρόκληση για την αποδοτική αντιμετώπιση και προσομοίωσή τους από ένα ρομπότ (Penrose R. , 1996).

Κεντρικό ρόλο στην κριτική του Penrose, έχει το στοιχείο της *κατανόησης* που απουσιάζει από κάθε καθαρά υπολογιστικό σύστημα, επιχείρημα που χρησιμοποιήθηκε και στη συλλογιστική του Gödel. Πέραν όμως αυτού του χαρακτηριστικού, οι άνθρωποι ως ενσυνειδήτα όντα διαθέτουν την ικανότητα της *άμεσης αίσθησης* των πραγμάτων. Η αδυναμία ενός ρομπότ να εκτιμήσει μια σονάτα του Μπετόβεν, μια έναστρη νύχτα όπως απεικονίζεται σε πίνακες του Van Gogh ή την αδιαμφισβήτητη ευφύια της συλλογιστικής στοχαστών, όπως ο Einstein, ίσως οφείλεται απλά στο γεγονός ότι «δεν μπορεί να αισθανθεί όλα όσα εμείς είμαστε σε θέση να αισθανθούμε όταν προσεγγίζουμε την παρουσία μιας τέτοιας ποιότητας» (Penrose R., 1996, p.399). Ο Penrose αναρωτιέται ωστόσο αν η παραγωγή ενός μεγαλειώδους έργου τέχνης βρίσκεται εντός των δυνατοτήτων ενός έξυπνα προγραμματισμένου υπολογιστή.

Για τον συγγραφέα κάτι τέτοιο δεν είναι εφικτό, από τη στιγμή που ένας υπολογιστής δεν διαθέτει τα αισθητηριακά χαρακτηριστικά ώστε να διακρίνει το ωραίο από το άσχημο, το εντυπωσιακό από το απλά ικανοποιητικό. Αλλά ένα τέτοιο σύστημα, δεν χρειάζεται να *αισθάνεται* για να αναπτύξει τα δικά του αισθητικά κριτήρια· θα μπορούσε έπειτα από μια περίοδο (bottom-up) εκπαίδευσης να αναπτύξει δικές του κρίσεις. Ακόμα κι εδώ όμως, «τα κριτήρια αυτά πρέπει να αποτελούν τμήμα των δεδομένων του υπολογιστή» (Penrose R., 1996, p.400), τα οποία θα εισάγονται μετά από μια λεπτομερή (top-down) ανάλυση κάποιων ανθρώπων ειδικών σε θέματα αισθητικής. Σήμερα υπάρχουν πλείστα παραδείγματα τέτοιων συστημάτων TN διαθέσιμα στο διαδίκτυο. Έχουν αναπτυχθεί αλγόριθμοι AI που μπορούν να δημιουργήσουν πίνακες, να συνθέσουν μουσική ή ακόμα να συγγράψουν ποίηση και πεζογραφία. Μερικά παραδείγματα προγραμμάτων αποτελούν το «DeepDream» (για δημιουργία εικόνων από επεξεργασμένες φωτογραφίες), το «Magenta» (για σύνθεση μουσικής), το «ChatGPT» (για συγγραφή ποίησης ή στίχων).

Ωστόσο η τέχνη - ειδικά τα πιο εκλεπτυσμένα δημιουργήματά της - συχνά περιλαμβάνει πολύπλοκα συναισθηματικά και διανοητικά στοιχεία, βαθιά συνδεδεμένα με την ανθρώπινη εμπειρία, τον πολιτισμό και την ιστορία· ιδιότητες μοναδικές στην ανθρώπινη συνείδηση. Για τον Penrose, το αισθητικό αποτέλεσμα τέτοιων προγραμμάτων δεν *εκφράζει* τίποτα, «επειδή ο ίδιος ο υπολογιστής δεν *αισθάνεται* τίποτα» (Penrose R., 1996, p.400).

Κι όμως, από την άλλη, η ανάπτυξη πανίσχυρων υπολογιστών που παίζουν σκάκι αποτελεί ένα τρανταχτό παράδειγμα στο οποίο οι υπολογιστές μπορούν να αποδειχθούν τρομερά αποτελεσματικοί. Το σκάκι είναι αναμφίβολα μια δραστηριότητα όπου η ισχύς της ανθρώπινης νόησης καθίσταται ιδιαίτερα φανερή. Πλέον είναι γνωστό ότι ένας υπολογιστής υπερέχει κατά πολύ έναντι του ανθρώπου στο συγκεκριμένο τομέα, ενώ ακόμα και οι πρωταθλητές του αθλήματος θα πρέπει να δώσουν τον καλύτερό τους εαυτό για να συναγωνιστούν τα καλύτερα υπολογιστικά συστήματα. Πέρα από αυτό το πεδίο, υπάρχουν αρκετοί ακόμα τομείς, όπως οι άμεσες αριθμητικές υπολογιστικές διαδικασίες, στους οποίους οι ικανότητες των υπολογιστών ξεπερνούν κατά πολύ αυτές του ανθρώπου.

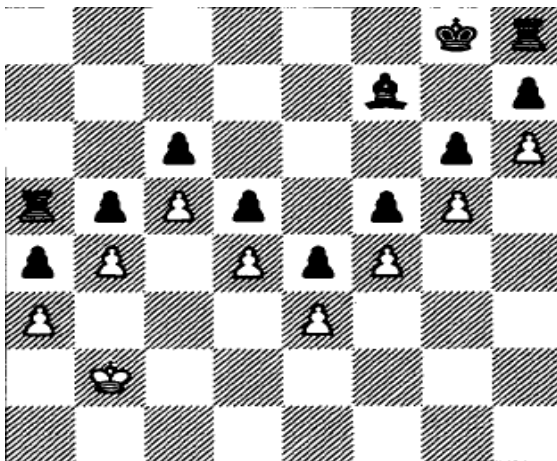
Μπορούμε ωστόσο να συμπεράνουμε ότι σε αυτές τις περιπτώσεις ο υπολογιστής επιτυγχάνει κάποια γνήσια κατανόηση του τι πραγματικά κάνει; Στην περίπτωση της

καθοδικής⁵ οργάνωσης, το σύστημα λειτουργεί με επιτυχία όχι λόγω της ουσιαστικής κατανόησης του τί κάνει, αλλά εξαιτίας της αποτελεσματικής χρήσης των γνώσεων των προγραμματιστών στη δημιουργία του προγράμματος.

Αντίθετα, σε περιπτώσεις όπου έχουμε ανοδική οργάνωση, δεν καθίσταται τόσο σαφής η απαίτηση για οποιαδήποτε συγκεκριμένη κατανόηση ως κανονιστική λειτουργία του συστήματος, τόσο για τη μηχανή TN όσο και για τον προγραμματιστή της. Εξαιρούνται από τα παραπάνω βέβαια, οι απαραίτητες ανθρώπινες γνώσεις για το σχεδιασμό των λεπτομερειών των αλγορίθμων για τη βελτίωση της απόδοσης του συστήματος και για την ίδια τη σύλληψη της ιδέας ότι ένα σύστημα μπορεί να βελτιώσει την απόδοσή του βασιζόμενο σε προηγούμενη εμπειρία. Το πρόβλημα που εμφανίζεται εδώ έχει να κάνει με τη σχετική ασάφεια του όρου *κατανόηση*, καθώς θα μπορούσε κάποιος να θεωρήσει ότι αυτά τα συστήματα διαθέτουν όντως κάποιο είδος κατανόησης, απλά με τους δικούς τους όρους.

Για την απόδειξη της απουσίας πραγματικής κατανόησης από τους σημερινούς υπολογιστές, ο Penrose παρουσιάζει μια παρτίδα σκάκι που τα πύονια της έχουν την παρακάτω διάταξη (Εικόνα 2).

Στη διάταξη αυτή, τα μαύρα πύονια φαίνεται να προηγούνται λόγω της ύπαρξης αξιωματικού και δύο πύργων στη σύνθεσή τους. Τα λευκά μπορούν, ωστόσο, να αποφύγουν την ήττα μετατοπίζοντας κατάλληλα το βασιλιά. Τα λευκά δε διατρέχουν κίνδυνο από τα μαύρα καθώς το φράγμα που έχουν στήσει είναι απόρθητο, κάτι που ένας άνθρωπος στοιχειωδώς εξοικειωμένος με τους κανόνες του σκακιού το αντιλαμβάνεται εύκολα. Η επόμενη κίνηση όμως, του Deep Thought, ενός πολύ ισχυρού υπολογιστή-παίκτη σκακιού, ήταν να πάρει το μαύρο πύργο με το πύονι του, σπάζοντας το φράγμα και οδηγούμενος σε αναπόφευκτη ήττα (Penrose R., 1996, p.46).



Εικόνα 2: Παρτίδα σκάκι

⁵ Μια υπολογιστική διαδικασία έχει καθοδική δομή, σύμφωνα με τον Penrose, αν έχει δημιουργηθεί σύμφωνα με κάποια ξεκάθαρα ορισμένη και απόλυτα κατανοητή υπολογιστική διαδικασία. Η διαδικασία αυτή οδηγεί σε μια ξεκάθαρη λύση ενός συγκεκριμένου προβλήματος, μέσω καλά ορισμένων κανόνων λειτουργίας και αποθήκευσης γνώσης, καθορισμένων εκ των προτέρων.

Κάποιοι υποστηρίζουν ότι για να μπορέσει ένα σύστημα ΤΝ να αποκτήσει κάποιου είδους πραγματική γνώση, θα πρέπει να είναι προγραμματισμένο ώστε να περιλαμβάνει ανοδικές διαδικασίες με έναν πολύ πιο στοιχειώδη τρόπο από ότι συνηθίζεται σήμερα στους υπολογιστές που παίζουν σκάκι. Έτσι, η τροφοδότηση του με ειδικούς καθοδικούς αλγοριθμικούς κανόνες δε θα ήταν απαραίτητη, καθώς οι γνώσεις του θα αναπτύσσονταν σταδιακά, με συσσώρευση εμπειριών. Οι καθοδικοί κανόνες, που είναι αρκετά απλοί και κατανοητοί για μας, σύμφωνα με τον Penrose, «δεν μπορούν από μόνοι τους να αποτελέσουν μια υπολογιστική βάση για πραγματική κατανόηση» (Penrose R., 1996, p.47). Αλλά ο συγγραφέας δε σταματά εδώ, καθώς θεωρεί ότι ούτε οι ανοδικές υπολογιστικές διαδικασίες θα μπορούσαν να παίξουν αυτόν τον ρόλο. Η γνήσια ανθρώπινη αντίληψη δεν μπορεί να υποκατασταθεί από τα σημερινά υπολογιστικά συστήματα, σε καμιά σημαντική περιοχή εξειδικευμένης γνώσης, όπου η γνήσια και συνεχής ανθρώπινη κατανόηση και διαίσθηση είναι απαραίτητες. Υποστηρικτές της θέσης Α και Β, ίσως θεωρήσουν ότι η εμφάνιση τέτοιων χαρακτηριστικών σε υπολογιστικά συστήματα είναι απλά ζήτημα χρόνου και τεχνολογικής προόδου στον τομέα της ΤΝ. Κι εδώ ο συγγραφέας επιμένει ότι θεμελιώδεις περιορισμοί εμφανίζονται σε κάθε καθαρά υπολογιστικό σύστημα, είτε καθοδικό είτε ανοδικό· η ψευδαίσθηση για την κατοχή *πραγματικής* γνώσης από τους υπολογιστές κάποια στιγμή διαλύεται (Penrose R., 1996, p.48).

Ο Penrose αποσκοπεί στο να δείξει ότι η μαθηματική κατανόηση συγκεκριμένα, δεν μπορεί να αναχθεί σε κάποια υπολογιστική διαδικασία. Κάτι τέτοιο θα φαινόταν προκλητικό στους οπαδούς της ΤΝ που θεωρούν την εκτέλεση αλγεβρικών και αριθμητικών υπολογισμών - ως προχωρημένες διεργασίες της ανθρώπινη εξέλιξης -, εκ των ων ουκ άνευ, προνομιακό πεδίο των υπολογιστών. Από την άλλη, ικανότητες που αναπτύχθηκαν νωρίς στην πορεία του ανθρώπινου είδους όπως π.χ. βάδισμα ή η αντίληψη πολύπλοκων οπτικών εικόνων, αποδεικνύεται αρκετά δύσκολο να γίνουν κτήμα μηχανών ελεγχόμενων από υπολογιστές, ενώ για τους ανθρώπους προκύπτουν αβίαστα. Για τον Penrose, για κάθε πολύπλοκη δραστηριότητα που μπορεί να περιλαμβάνει την πραγματοποίηση μαθηματικών υπολογισμών ή μια παρτίδα σκάκι ή καθημερινές συνηθισμένες πράξεις μας, οι υπολογιστές μπορεί να είναι εξαιρετικά αποτελεσματικοί από τη στιγμή που η παραπάνω δραστηριότητα έχει γίνει κατανοητή με βάση κάποιους ξεκάθαρους υπολογιστικούς κανόνες.

3.6 Η νέα φυσική που χρειαζόμαστε για την εξήγηση της νόησης

Η «Ενορχηστρωμένη Θεωρία της Συνείδησης» (Orchestrated Objective Reduction), που προτάθηκε από τον Penrose, εστιάζει στις κβαντικές διεργασίες που συμβαίνουν εντός των μικροσωληνίσκων, δομικών συστατικών που βρίσκονται στους νευρώνες του εγκεφάλου.

Σύμφωνα με τη θεωρία Orch-OR, οι μικροσωληνίσκοι παίζουν καθοριστικό ρόλο στη δημιουργία της συνείδησης. Οι μικροσωληνίσκοι είναι κυλινδρικές πρωτεϊνικές δομές που σχηματίζουν ένα δίκτυο μέσα στους νευρώνες και εμπλέκονται σε διάφορες κυτταρικές λειτουργίες. Κβαντικά φαινόμενα, όπως η υπέρθεση και η διεμπλοκή, συμβαίνουν μέσα σε

αυτούς τους μικροσωληνίσκους, τα οποία είναι υπεύθυνα για τη συνειδητή εμπειρία (Penrose & Hameroff, 1996). Τέτοια γεγονότα κβαντικής φύσεως εντός των μικροσωληνίσκων επηρεάζουν τη δυναμική της νευρικής δραστηριότητας και συμβάλλουν στην εμφάνιση της ενσυνείδητης σκέψης. Συγκεκριμένα, η κατάρρευση της κβαντικής υπέρθεσης, μια διαδικασία γνωστή ως αντικειμενική αναγωγή (OR), προτείνεται ως ο μηχανισμός σύμφωνα με τον οποίο προκύπτουν οι στιγμές συνειδητότητας. Αυτές οι ενορχηστρωμένες αντικειμενικές αναγωγές συμβαίνουν με μη αλγοριθμικό τρόπο, παρέχοντας μια μη υπολογιστική βάση για τη συνείδηση. Σύμφωνα με τη θεωρία Orch-OR, οι κβαντικές διεργασίες που συμβαίνουν στους μικροσωληνίσκους προστατεύονται από την αποσυγκόλληση, δηλαδή την απώλεια της κβαντικής συνοχής που προκαλείται από τις αλληλεπιδράσεις με το περιβάλλον. Η προστασία αυτή οφείλεται στις μοναδικές ιδιότητες των πρωτεϊνών των μικροσωληνίσκων, ιδίως στη διατεταγμένη δομή του πλέγματος και στο υδρόφοβο εσωτερικό τους. Αυτά τα χαρακτηριστικά δημιουργούν ένα περιβάλλον που επιτρέπει τη διατήρηση και τη διάδοση των κβαντικών καταστάσεων, επιτρέποντας τη συνοχή που είναι απαραίτητη για τη συνείδηση. Αυτό συμβαίνει διαρκώς στον εγκέφαλο. Ίσως δεν χρειάζεται τελικά να κατασκευάσουμε έναν κβαντικό υπολογιστή, αφού όλοι μεταφέρουμε έναν μέσα στο κεφάλι μας (Penrose & Hameroff, 1996).

Οι επικριτές της θεωρίας Orch-OR έχουν διατυπώσει διάφορες αντιρρήσεις τόσο σε σχέση με την επιστημονική της βάση όσο και με τις φιλοσοφικές της συνέπειες. Μια από τις σημαντικότερες κριτικές της θεωρίας στέκεται στην έλλειψη εμπειρικών αποδείξεων για την ύπαρξη κβαντικών διεργασιών στις σχετικές κλίμακες. Ακόμα, αμφισβητείται η δυνατότητα φαινομένων κβαντικού επιπέδου να παίζουν σημαντικό ρόλο σε ένα μακροσκοπικό βιολογικό σύστημα όπως ο εγκέφαλος. Ορισμένοι υποστηρίζουν ότι οι προτεινόμενες κβαντικές διεργασίες εντός των μικροσωληνίσκων είναι πολύ εύθραυστες για να διατηρήσουν την συνοχή που είναι απαραίτητη για την παραγωγή συνειδητής εμπειρίας. Επιπλέον, υπάρχει συνεχής συζήτηση σχετικά με την πραγματική φύση των κβαντικών φαινομένων στα βιολογικά συστήματα και κατά πόσον είναι σημαντικά για την κατανόηση της συνείδησης.

Παρά τις αντιρρήσεις που την περιβάλλουν, η θεωρία των Penrose-Hameroff έχει προκαλέσει σημαντικό ενδιαφέρον και συζήτηση στον τομέα των μελετών της συνείδησης, καθώς προσφέρει μια προοπτική που ενσωματώνει την κβαντική φυσική στην κατανόηση της συνείδησης, αμφισβητώντας τις πιο συμβατικές, κλασικές προσεγγίσεις. Μελλοντικές έρευνες στην κατεύθυνση της θεωρίας μπορεί να περιλαμβάνουν τη διερεύνηση της προθετικότητας (intentionality) και των μηχανισμών των κβαντικών διεργασιών στους μικροσωληνίσκους, τη διεξαγωγή πειραματικών μελετών για τον έλεγχο των προβλέψεων της θεωρίας και την περαιτέρω διερεύνηση των επιπτώσεων των φαινομένων κβαντικού επιπέδου στην κατανόηση της συνείδησης. Αν και η θεωρία Orch-OR παραμένει αμφιλεγόμενη, ανοίγει νέους δρόμους έρευνας σχετικά με τη φύση της συνείδησης και την πιθανή αλληλεπίδραση μεταξύ κβαντικής φυσικής και νευροεπιστήμης.

3.7 Κριτική και αξιολόγηση των θέσεων του Penrose

Ενώ οι ιδέες του Penrose συγκέντρωσαν την προσοχή των μελετητών και επιστημόνων, πυροδοτώντας συζητήσεις στα σχετικά πεδία προβληματισμού, έχουν επίσης αντιμετωπιστεί με κριτική και αμφιβολίες από διάφορες θεωρητικές σκοπιές. Αρκετά βασικά σημεία κριτικής και αξιολόγησης μπορούν να παρατεθούν σχετικά με τις απόψεις του Penrose για τη φύση των ενσυνείδητων πράξεων.

Η θεωρία του βασίζεται στην ιδέα ότι η κβαντική συνοχή μπορεί να διατηρηθεί και να χρησιμοποιηθεί σε ένα βιολογικό σύστημα όπως ο εγκέφαλος, παρά τις δυσκολίες της διατήρησης κβαντικών καταστάσεων σε ένα θερμό και θορυβώδες περιβάλλον. Οι σκεπτικιστές υποστηρίζουν ότι η εν λόγω θεωρία χρειάζεται περισσότερα εμπειρικά στοιχεία και ελέγξιμες προβλέψεις για να κερδίσει ευρύτερη αποδοχή από την επιστημονική κοινότητα. Ακόμα, πολλοί επικριτές υποστηρίζουν ότι τα κβαντικά φαινόμενα, ακόμα κι αν λαμβάνουν χώρα στον εγκέφαλο, μπορεί να μην είναι άμεσα υπεύθυνα για τις συνειδητές εμπειρίες. Υποστηρίζουν ότι οι πολύπλοκες διεργασίες της συνείδησης είναι πιο πιθανό να προκύπτουν από αναδυόμενες ιδιότητες ή διαδικασίες υψηλότερου επιπέδου στον εγκέφαλο, παρά από κβαντικά φαινόμενα. Η άποψη αυτή ευθυγραμμίζεται με τις επικρατούσες προσεγγίσεις που εστιάζουν στη θεωρία των νευρωνικών δικτύων και την κατανομημένη επεξεργασία πληροφοριών ως τη βάση της συνείδησης (βλ. Tononi, Koch, Churchland).

Επιπλέον, η προοπτική του Penrose έχει επικριθεί για τις δυϊστικές της προεκτάσεις. Ορισμένοι υποστηρίζουν ότι η εστίασή του σε μη υπολογιστικές, κβαντικές διεργασίες ως πηγή της συνείδησης διαιώνίζει τον δυισμό νου-σώματος, ο οποίος προϋποθέτει έναν διαχωρισμό μεταξύ του φυσικού εγκεφάλου και του άυλου νου. Αυτή η δυϊστική στάση έρχεται σε αντίθεση με μια πιο μονιστική ή υλιστική προοπτική που επιδιώκει να εξηγήσει τη συνείδηση αποκλειστικά με όρους φυσικών διεργασιών στον εγκέφαλο. Οι επικριτές υποστηρίζουν ότι η προσέγγιση του Penrose δεν διαθέτει μια σαφή συσχέτιση μεταξύ του φυσικού και του νοητικού, εγείροντας ενστάσεις σχετικά με τη συνολική συνοχή της θεωρίας του.

3.7.1 Η κριτική του Paul Churchland

Μπορούν τα είδη των αλγοριθμικών διαδικασιών που εκτελούνται από ένα τυπικό υπολογιστικό πρόγραμμα να καλύψουν το πλήρες φάσμα της ανθρώπινης μαθηματικής γνώσης και ικανότητας; Η αρνητική απάντηση που θα δώσει ο Penrose βρίσκει σύμφωνο τον Churchland, αν και για διαφορετικούς λόγους. Σύμφωνα με τον Churchland, ο Penrose όπως και πολλοί άλλοι, επικαλείται το θεώρημα του Gödel σχετικά με τη μη πληρότητα κάθε αξιωματικοποίησης των μαθηματικών. Όπως αναφέρθηκε, το περίφημο θεώρημα αποδεικνύει ότι κανένα πεπερασμένο σύνολο αλγοριθμικών διαδικασιών δεν μπορεί να παραγάγει όλες τις μαθηματικές αλήθειες. Πρέπει πάντα να υπάρχουν κάποιες μαθηματικές αλήθειες, που αποδεικνύονται εύκολα με τρόπους που βρίσκονται έξω από το εν προκειμένω συγκεκριμένο αλγοριθμικό σύστημα, αλλά δεν δύναται να αποδειχτούν εντός

του. Καμιά μηχανική υλοποίηση, λοιπόν, αυτού του αλγοριθμικού συστήματος δεν μπορεί να παραγάγει όλες τις μαθηματικές αλήθειες που εμείς, οι άνθρωποι, μπορούμε να αποδείξουμε. Ο Penrose, όπως πολλοί άλλοι, ερμηνεύει αυτό το γεγονός ως απόδειξη τού ότι η ανθρώπινη γνώση των μαθηματικών αληθειών δεν μπορεί να εξηγηθεί πλήρως από την υπόθεση ότι χρησιμοποιούμε αλγοριθμικές διαδικασίες (Churchland, 1996, p.246).

Ο Churchland επισημαίνει, ωστόσο, ότι η ερμηνεία του Penrose για το θεώρημα του Gödel δεν συμφωνεί με την πλειοψηφία των ερμηνειών άλλων στοχαστών. Η συνήθης απάντηση είναι ότι, εφόσον ο Gödel είχε δίκιο, πρέπει κι εμείς οι άνθρωποι να έχουμε ανάλογους περιορισμούς. Πρέπει δηλαδή να υπάρχουν μαθηματικές αλήθειες που βρίσκονται πέρα από το δικό μας οπλοστάσιο αλγοριθμικών διεργασιών, αλήθειες τις οποίες ένα ανώτερο ον με μεγαλύτερο οπλοστάσιο θα μπορούσε να αποδείξει. Για τον Churchland δεν αποτελεί έκπληξη το γεγονός ότι οι άνθρωποι μπορούν να αποδείξουν περισσότερα από μια μηχανή η οποία περιορίζεται στη χρήση των κλασικών αξιωμάτων Peano και ορισμένων κανόνων. «Το αποτέλεσμα του Gödel δείχνει ότι, όσον αφορά τη μαθηματική γνώση, οι άνθρωποι δεν περιορίζονται σε ένα μόνο σύνολο αλγοριθμικών διαδικασιών. Αλλά αυτό είναι σύμφωνο με την υπόθεση ότι κάποιου είδους αλγοριθμικές διεργασίες αποτελούν το υπόβαθρο της διάσπαρτης μαθηματικής μας γνώσης, όποια κι αν είναι αυτή» (Churchland, 1996, p.246).

Η επιθυμία του Penrose να υπερασπιστεί την υπόθεση ότι οι άνθρωποι κατέχουν μια μη αλγοριθμική ικανότητα που τους επιτρέπει να αναγνωρίζουν μαθηματικές αλήθειες, συγκλίνει με αυτή του Churchland. Ακόμα, ο Churchland φαίνεται να αποδέχεται «μια μορφή ‘ενόρασης’ η οποία δεν εξαρτάται από το χειρισμό φυσικών συμβόλων, ένα χειρισμό που καθοδηγείται από κανόνες, σε μια διεργασία που αποτελείται από διακριτές καταστάσεις» (Churchland, 1996, p.247). Εδώ, πρέπει να επισημανθεί η διαφωνία του Churchland για το που βρίσκεται η εν λόγω μη αλγοριθμική ικανότητα.

Ο Penrose στην απόπειρά του να περιγράψει αυτήν την ικανότητα, εισάγει τον αναγνώστη στο πεδίο της κβαντικής φυσικής και πιο συγκεκριμένα, στο πεδίο των κβαντικών βαρυτικών επιδράσεων που συμβαίνουν εντός του εγκεφάλου μας. Εκεί αναζητά και εντοπίζει την παραπάνω ικανότητα, παρά το γεγονός της υποθετικής και προς το παρόν αναπόδεικτης ύπαρξης του τελευταίου πεδίου. Ισχυρίζεται ακόμα, ότι τέτοιες διεργασίες (π.χ. ιδιότητες της υπέρθεσης και κατάρρευσης των κυμάτων) μπορεί να ενσωματώνουν μη κλασικούς υπολογισμούς, υπολογισμούς δηλαδή, απαραίτητους για την επεξήγηση του φαινομένου της ‘ενόρασης’ στο πεδίο της ανθρώπινης μαθηματικής γνώσης. Έτσι κι αλλιώς, όλοι συνηγορούν στο ότι η διαδικασία της κυματικής υπέρθεσης και της συνακόλουθης κατάρρευσης προς μια κλασική κατάσταση αποτελεί μια μη αλγοριθμική διαδικασία.

Τη στιγμή που δέχεται τη μη αλγοριθμική φύση των διαδικασιών, ταυτόχρονα ο Churchland αντικρούει τη θέση του Penrose τονίζοντας ότι «δε φαίνεται να υπάρχει κάτι συγκεκριμένο που να υποδηλώνει ότι οι κβαντικές διεργασίες πραγματοποιούν αναγνωρίσιμους υπολογισμούς» (Churchland, 1996, p.247). Δεν υπάρχει καμία αναγνωρίσιμη οδός μέσω της οποίας πληροφορίες σχετικές με κάποια μακροσκοπικά

μαθηματικά προβλήματα - «μια πολύπλοκη τετραγωνική εξίσωση που βρίσκεται μπροστά στα μάτια μου, για παράδειγμα» (Churchland, 1996, p.246) - θα μπορούσαν να φτάσουν μέχρι το επίπεδο των υπερτιθέμενων κβαντικών καταστάσεων, περνώντας μέσα από τον υπολογιστικό διάυλο της κατάρρευσης του κύματος, και μετά πίσω πάλι στο κλασικό επίπεδο, με έναν τρόπο που θα βοηθούσε στην αναγνώριση της περίπλοκης εξίσωσης ως μιας ακόμη περίπτωσης της οικείας τετραγωνικής μορφής. Ένα επιπλέον επιχείρημα που χρησιμοποιεί ο Churchland εναντίον της θέσης του Penrose στηρίζεται στη διαφορά κλίμακας. Οι διεργασίες του εγκεφάλου που μπορούμε πραγματικά να εμπλέξουμε στο έργο της επεξεργασίας πληροφοριών είναι όλες στην κλίμακα των ανταλλαγών μάζας-ενέργειας, η οποία βρίσκεται πολύ πιο πάνω από το κβαντικό επίπεδο και ξεκάθαρα μέσα στο κλασικό.

Ο Churchland ολοκληρώνει την κριτική του στον Penrose κάνοντας λόγο για πληθώρα μη αλγοριθμικών διεργασιών πριν και εκτός του κβαντικού επιπέδου. Φέρνοντας ως παράδειγμα τις διεργασίες που πραγματοποιούνται στο hardware ενός νευρωνικού δικτύου, οι οποίες φέρουν και το κύριο βάρος της εγκεφαλικής υπολογιστικής δραστηριότητας, ξεκαθαρίζει ότι «είναι μη αλγοριθμικές, με την έννοια ότι δε συνίστανται σε μια σειρά διακριτών φυσικών καταστάσεων που πραγματώνονται εν σειρά, υπό την καθοδήγηση ενός αποθηκευμένου συνόλου κανόνων για το χειρισμό συμβόλων» (Churchland, 1996, p.248). Αντ' αυτού αποτελούν αναλογικές διεργασίες, τα στοιχεία και οι δραστηριότητες των οποίων, παίρνουν πραγματικές τιμές ενώ εκτυλίσσονται παράλληλα και σύμφωνα με τους φυσικούς νόμους.

Κεφάλαιο 4 : Η φιλοσοφική ματιά του Paul Churchland στα ερωτήματα της συνείδησης

4.1 Εισαγωγή στη σκέψη του P.Churchland

«Ξαφνικά κοιτάζουμε σε έναν καθρέφτη. Όχι τους μακρινούς ουρανοί ούτε το βιολογικό μας παρελθόν, ούτε ακόμη τον απέραντο μικρόκοσμο, αλλά τον εαυτό μας. Είναι η βασική μας αντίληψη για την ανθρώπινη νόηση και δράση ένας ακόμη μύθος, χρήσιμος ίσως σε κάποιο βαθμό στο παρελθόν, αλλά λανθασμένος στην επιφάνεια ή στον πυρήνα του; Μήπως μια κατάλληλη θεωρία της εγκεφαλικής λειτουργίας θα παρουσιάσει μια πολύ διαφορετική ή ασύμβατη εικόνα της ανθρώπινης φύσης; Θα έπρεπε να προετοιμάσουμε συναισθηματικά τον εαυτό μας για μια ακόμη εννοιολογική επανάσταση που θα μας αγγίξει περισσότερο από οποιαδήποτε άλλη;» (Churchland, 1996, p.18).

Το βιβλίο «Η μηχανή της λογικής και η θέση της ψυχής» (The Engine of Reason, The Seat of the Soul) παρέχει μια ολοκληρωμένη επισκόπηση της φιλοσοφίας και των ιδεών του Churchland σχετικά με τη φύση του νου και της συνείδησης. Το βιβλίο παρουσιάζει μια υλιστική προσέγγιση για την κατανόηση του νου, υποστηρίζοντας ότι τα νοητικά φαινόμενα μπορούν να εξηγηθούν από τις φυσικές διεργασίες που συμβαίνουν στον εγκέφαλο. Ο Churchland αμφισβητεί τις δυϊστικές απόψεις για τον νου, υποστηρίζοντας την επιστημονική και εμπειρική κατανόηση της συνείδησης. Στην ανάπτυξη της επιχειρηματολογίας του, ο Churchland διερευνά μια ποικιλία θεμάτων που κυμαίνονται από την αισθητηριακή αναπαράσταση, την παράλληλη κατανομημένη επεξεργασία και τη διανυσματική κωδικοποίηση μέχρι το ρόλο που παίζει η μνήμη και η γλώσσα στις γνωσιακές μας διαδικασίες. Βασίζεται σε ευρήματα από τη νευροεπιστήμη, τη γνωστική ψυχολογία και την τεχνητή νοημοσύνη για να υποστηρίξει την υλιστική του θέση και εξετάζει πώς ο εγκέφαλος επεξεργάζεται πληροφορίες, κατασκευάζει αναπαραστάσεις του κόσμου και παράγει συνειδητές εμπειρίες, δίνοντας ιδιαίτερη έμφαση στη μελέτη των διαδικασιών που συμβαίνουν εντός των επανατροφοδοτικών δικτύων.

4.2 Η κριτική του Churchland στο Turing τεστ

Ο Alan Turing (1912-1954) ήταν ένας Βρετανός μαθηματικός και επιστήμονας των υπολογιστών. Θεωρείται ένας από τους πιο επιδραστικούς επιστήμονες του 20ού αιώνα στους τομείς των υπολογιστών, της λογικής, και της τεχνητής νοημοσύνης.

Μια από τις πιο σημαντικές συνεισφορές αυτού του ιδιοφυούς επιστήμονα, αποτελεί ίσως το περίφημο Turing Test. Το τεστ έχει ως εξής. Μια υποψήφια υπολογιστική μηχανή και ένας άνθρωπος βρίσκονται σε ξεχωριστά δωμάτια, μακριά και χωρίς δυνατότητα επικοινωνίας με τους κριτές. Οι κριτές μπορούν μέσω μιας αμφίδρομης τηλετυπικής συσκευής να επικοινωνούν τόσο με τη μηχανή όσο και με τον άνθρωπο. Αυτή η πληροφοριακή οδός είναι το μόνο μέσο με το οποίο μπορούν οι κριτές να πληροφορηθούν για τους δύο υποψήφιους. Χωρίς να έχουν πρόσβαση στον τόνο της φωνής, στη γλώσσα του

σώματος και τα υπόλοιπα φυσικά χαρακτηριστικά, οι κριτές πρέπει να ξεχωρίσουν τον άνθρωπο από τη μηχανή στηριζόμενοι αποκλειστικά σε ερωταπαντήσεις με τους υποψηφίους μέσω της συσκευής. Πιθανά θέματα ερωτήσεων θα έχουν να κάνουν με το συναισθηματικό κόσμο των υποψηφίων, τις πολιτικές απόψεις, τις κοινωνικές τους δεξιότητες ανάμεσα σε άλλα, στην προσπάθεια να βρεθεί εκείνη η πτυχή της ανθρώπινης συμπεριφοράς στην οποία η μηχανή υπολείπεται. Σύμφωνα με τον Turing, η αποτυχία των κριτών να διακρίνουν τον άνθρωπο από τη μηχανή, θα μας ανάγκαζε να αποδεχθούμε ότι δεν υπάρχει κάποιο μοναδικό χαρακτηριστικό στον άνθρωπο - η συνείδηση συγκεκριμένα - το οποίο δεν μπορεί να κατέχει και να αναπτύξει, εξίσου καλά, μια μηχανή.

Παρουσιάζοντας διαφορετικές και βελτιωμένες παραλλαγές στο πρωτότυπο πείραμα, ο Churchland αμφιβάλει για την αξιοπιστία του τεστ. Το πρώτο συμπέρασμα όπου καταλήγει, εξετάζοντας τα διάφορα πειράματα, είναι ότι η βιομηχανία παραγωγής προγραμμάτων Τεχνητής Νοημοσύνης – παρά τα επιτεύγματά της – απέχει, ακόμα, αρκετά από το να δημιουργήσει κάποιο λειτουργικό σύστημα που να προσεγγίζει την ανθρώπινη νόηση. Σαν δεύτερο συμπέρασμα σημειώνει ότι ακόμα και οι άνθρωποι κρίνονται αρκετές φορές αναξιόπιστοι στο να διακρίνουν την ανθρώπινη νοημοσύνη από τις μηχανικές προσομοιώσεις της (Churchland, 1996, pp.228-233).

Οι περισσότερες κριτικές του Turing τεστ εστιάζουν στην επιείκιά του, καθώς ενδέχεται να αποδώσει αυθεντικές νοητικές ικανότητες σε μη νοήμονα όντα. Μια πειστική λεκτική συμπεριφορά διαμέσου τηλετυπικής επικοινωνίας θα μπορούσε να οφείλεται σε διάφορες αιτίες, αιτίες μη σχετιζόμενες ουσιαστικά με την αληθινή ενσυνείδητη νοημοσύνη. Προτείνεται ως απάντηση, να συμπεριληφθεί μια ευρύτερη ποικιλία συμπεριφορών με στόχο την αυστηροποίηση των κριτηρίων. Αλλά πώς θα διευρυνθούν τα κριτήρια χωρίς να είναι ξεκάθαρο ποιοι τρόποι συμπεριφοράς σχετίζονται με την κατοχή ενσυνείδητης νοημοσύνης; Από την άλλη, εμφανίζονται κριτικές που τονίζουν την περιοριστικότητα του τεστ σε σχέση με την κατοχή της γλωσσικής ικανότητας. Νοήμονα όντα, όπως παιδιά πριν την ηλικία της ομιλίας, ενήλικες με τοπική αφασία, τα περισσότερα ανώτερα ζώα του πλανήτη είναι καταδικασμένα να αποτύχουν στο τεστ, καθώς δεν επικοινωνούν μέσω της ανθρώπινης γλώσσας. «Ίσως λοιπόν», σημειώνει ο Churchland, «δε θα έπρεπε να υποθέσουμε ότι το τεστ Turing επιβάλλει κατ'αρχήν μια αναγκαία συνθήκη για την ύπαρξη ενσυνείδητης νοημοσύνης» (Churchland, 1996, p.234).

Ανεξάρτητα από τα πλεονεκτήματα και τα μειονεκτήματα του τεστ, οι ερευνητές αναγκαστικά επιστρέφουν στη «συμπεριφορική ομοιότητα με την παραδειγματική περίπτωση (τον άνθρωπο)», στο βαθμό που απουσιάζει μια επαρκής θεωρία του τί είναι νοημοσύνη και πώς υλοποιείται στα φυσικά συστήματα. Μια τέτοια ολοκληρωμένη θεωρία θα απέφευγε τις συμπεριφορικές παγίδες και περιορισμούς του τεστ Turing, φανερώνοντας από μόνη της τις σχετικές ιδιότητες, συμπεριφορές, τεχνικές και μηχανισμούς που χαρακτηρίζουν την αυθεντική νοημοσύνη (Churchland, 1996, p.235).

Ο Churchland αναφέρει χαρακτηριστικά ότι «το τεστ Turing είναι ένα τεστ για ανθρώπους που δεν έχουν μια επαρκή θεωρία του τί είναι η νοημοσύνη ή καμία θεωρία πέρα από το συνηθισμένο πλαίσιο των προεπιστημονικών αντιλήψεών μας», εντάσσοντας

τον ίδιο τον Turing σε αυτό το πλαίσιο. Η φιλοδοξία του έγκειται παρ'όλ'αυτά στο ότι σήμερα θα μπορούσαμε να ελέγξουμε απευθείας την ύπαρξη αυτών των χαρακτηριστικών, κοιτώντας στο εσωτερικό του υποψήφιου συστήματος, για να δούμε τι συμβαίνει εντός του, ή εξετάζοντας τη συμπεριφορά του συστήματος σε πολύ ευρύτερα και απαιτητικότερα πλαίσια από αυτό του τηλετυπικού διαλόγου.

Σήμερα, η παραδειγματική περίπτωση του ανθρώπου είναι λιγότερο αινιγματική σε σχέση με το 1950, εφόσον οι εσωτερικές δομές και οι δραστηριότητές της δεν είναι πλέον τόσο απρόσιτες στην επιστημονική παρατήρηση (Churchland, 1996, p.235). Η αξιοσημείωτη πρόοδος των διάφορων - σχετικών με την ανθρώπινη κατανόηση - επιστημονικών κλάδων, έχει θέσει τα εννοιολογικά και πειραματικά θεμέλια, ενδεχομένως επαρκή για τη διατύπωση μιας ορθής θεωρίας για την ανθρώπινη, αλλά και ζωική νοημοσύνη. Κατ'αρχάς, η θεωρία πρέπει να είναι σε θέση να εξηγήσει το ευρύ φάσμα των συμπεριφορών που αντιλαμβάνεται και επιδεικνύει κάθε πραγματικό ζώο και όχι απλά την ικανότητα για έναν τηλετυπικό διάλογο. Σημασία πρέπει να δοθεί, επίσης, στη συνέπεια της θεωρίας ως προς τις εσωτερικές υπολογιστικές ικανότητες του συστήματος που παράγει τη συμπεριφορά· πρέπει, με άλλα λόγια, να συναρτάται κατάλληλα με τα κινηματικά και δυναμικά χαρακτηριστικά του βιολογικού εγκεφάλου. Οφείλει, εν τέλει, να παρέχει μια εξήγηση για τα θεμελιώδη χαρακτηριστικά του γινώσκουν, όπως η μάθηση, η αντιληπτική αναγνώριση και η εννοιολογική αλλαγή.

Με μια τέτοιου είδους θεωρία, ισχυρίζεται ο Churchland, η επιστημονική κοινότητα θα μπορούσε να απαντήσει πιο αποφασιστικά στο ερώτημα για τη νοημοσύνη των φυσικών ή τεχνητών συστημάτων. Η προσέγγιση του Churchland διαχωρίζεται, συνεπώς, αισθητά από αυτή του Turing, καθώς η πρώτη «τοποθετεί τις πολύπλοκες αιτίες της νοήμονος συμπεριφοράς στο κέντρο της προσοχής» πέρα από την παρατηρήσιμη συμπεριφορά που παράγουν αυτές οι αιτίες (Churchland, 1996, p.236).

4.3 Η υπολογιστική θεωρία της συνείδησης

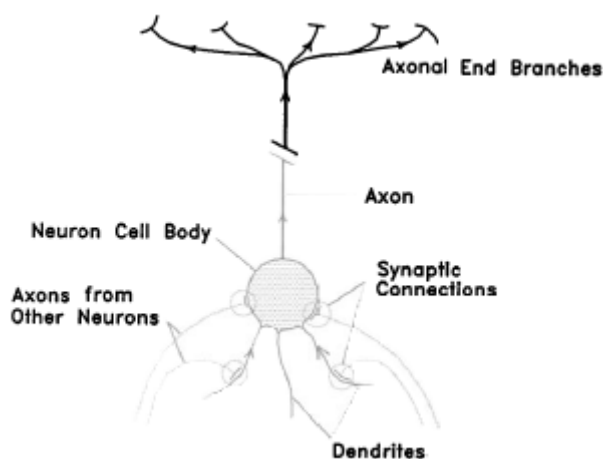
Η υπολογιστική θεωρία του Paul Churchland υποστηρίζει ότι η συνείδηση και οι γνωστικές διαδικασίες μπορούν να κατανοηθούν με όρους *παράλληλης καταναεμημένης επεξεργασίας*⁶ (*parallel distributed processing*) και νευρωνικών αναπαραστάσεων. Σύμφωνα με αυτή, ο νους θεωρείται ως ένα σύστημα επεξεργασίας πληροφοριών, παρόμοιο με έναν τεχνητό νευρωνικό δίκτυο.

Συγκεκριμένα, η ενσυνείδητη νοημοσύνη που επιδεικνύουν οι άνθρωποι επιτυγχάνεται μέσα σε δίκτυα μέσω διανυσματικών κωδικοποιήσεων και επεξεργασίας διανυσμάτων. Τέτοια δίκτυα αποτελούν τα δίκτυα επανατροφοδότησης και τα συστήματα τέτοιων δικτύων. Οι διάφορες μορφές των γνωσιακών μας ικανοτήτων αποκτώνται μέσω

⁶ Πρόκειται για το είδος υπολογισμού σύμφωνα με το οποίο ένα πρότυπο ή μοτίβο (pattern) μετασχηματίζεται σε ένα άλλο κατά τη διέλευσή του μέσα από μια μεγάλη διάταξη συναπτικών συνδέσεων. Χαρακτηριστικό κοινό σε όλο το ζωικό βασίλειο, με σημαντικά πλεονεκτήματα σε σχέση με την πιο γνωστή, *σειραϊκή επεξεργασία* (*serial processing*), η οποία χαρακτηρίζει τους κλασικούς υπολογιστές.

μιας διαδικασίας συμμόρφωσης των συναπτικών βαρών που κατανέμουν τους διάφορους χώρους των νευρωνικών ενεργοποιήσεων, σε ιεραρχίες πρωτοτυπικών κατηγοριών και ακολουθιών, σε ένα εννοιολογικό πλαίσιο που αποκρίνεται στις αντιληπτικές εισροές, επιτρέπει την εκούσια εξερεύνηση και κατευθύνει τη συμπεριφορική εκροή.

4.3.1 Ο ρόλος των νευρωνικών δικτύων στην ανθρώπινη νοημοσύνη



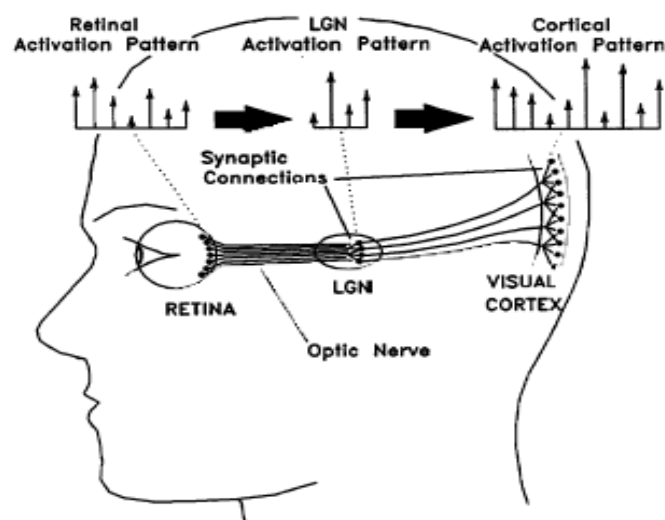
Εικόνα 3: Ένας τυπικός νευρώνας. Λαμβάνει διεγερτικά και ανασταλτικά σήματα από άλλους νευρώνες διαμέσου των πολλών συναπτικών συνδέσεων που σχηματίζουν στο κυτταρικό σώμα του νευρώνα και διαμέσου του εκτεταμένου δέντρου δενδριτικών κλάδων. Αθροίζει τα διάφορα εισερχόμενα σήματα και εκπέμπει το κατάλληλο σήμα κατά μήκος του νευράξονά του, για να αποκτήσει επαφή με άλλους νευρώνες (Churchland, 1994, p.4).

Τα γενικά χαρακτηριστικά του κόσμου που μένουν σταθερά και αναλλοίωτα, αναπαρίστανται στον εγκέφαλο από σχετικά αμετάβλητες διατάξεις συναπτικών συνδέσεων. Η αντίδραση του εγκεφάλου στις καθημερινές προκλήσεις που παρουσιάζονται - εύρεση τροφής, προστασία από εισβολείς, ανατροφή παιδιών κοκ. - καθορίζεται από διατάξεις προσεκτικά προσαρμοσμένων συνδέσεων. «Τέτοιες τυποποιημένες καταστάσεις έχουν λιγότερο ή περισσότερο τυποποιημένα αιτιώδη χαρακτηριστικά και απαιτούν τυποποιημένους – αλλά κατάλληλα πλαστικούς – τρόπους αντίληψης και συμπεριφορικής απόκρισης» (Churchland, 1994, p.5).

Απαραίτητη προϋπόθεση για την ανάπτυξη τέτοιων ικανοτήτων απόκρισης είναι η μάθηση της γενικής αιτιακής δομής του κόσμου, διαδικασία που λαμβάνει χώρα, κυρίως, κατά την βρεφική και παιδική ηλικία. Σε αυτή την περίοδο, οι ισχύες των συνδέσεων ή *βάρη* λαμβάνουν σταδιακά όλο και πιο χρήσιμες τιμές. Η μάθηση επιτυγχάνεται συχνά μέσω μιας διαδικασίας που ονομάζεται οπισθοδιάδοση, όπου το δίκτυο προσαρμόζει τα εσωτερικά του βάρη για να ελαχιστοποιήσει τα σφάλματα μεταξύ της εξόδου του και της επιθυμητής εξόδου-στόχου. Οι παράγοντες που συντελούν σε αυτήν τη διαδικασία αντανακλούν εν μέρει τις γενετικές καταβολές του ατόμου, αλλά περισσότερο τις ξεχωριστές εμπειρίες που συσσωρεύονται μέσα από την ανατροφή του παιδιού (Churchland, 1994, p.5).

Ωστόσο, μια σημαντικά διαφορετική διαδικασία αναπαράστασης λαμβάνει χώρα, όταν πρόκειται για τα επιμέρους και εφήμερα χαρακτηριστικά του άμεσου αισθητηριακού περιβάλλοντος του εγκεφάλου και τις συνεχείς εμπειρίες του. Αυτές οι στιγμιαίες μεταβολές «αναπαρίστανται από μια στιγμιαία διάταξη βαθμών ενεργοποίησης των πολλών νευρώνων του εγκεφάλου, όπως αυτών που βρίσκονται στον αμφιβληστροειδή και τον οπτικό φλοιό» (Churchland, 1994, p.6). Ενώ η μεταβολή των συναπτικών συνδέσεων των νευρώνων είναι μια αργή διαδικασία, δεν ισχύει το ίδιο για το βαθμό των εσωτερικών ενεργοποιήσεών τους, τον οποίο οι νευρώνες μπορούν να αλλάζουν στιγμιαία. Αυτός ο βαθμός ενεργοποίησης κάθε νευρώνα αναπροσαρμόζεται διαρκώς ως αποτέλεσμα των διεγέρσεων ή των αναστολών που πηγάζουν από τον εξωτερικό κόσμο. Έτσι, η κατάσταση του εγκεφάλου κάθε δεδομένη στιγμή προκύπτει από το συνολικό πρότυπο αυτών των βαθμών ενεργοποίησης. Όμοια, η χρονική ακολουθία αυτών των μεταβαλλόμενων προτύπων σχηματίζει, εντός του εγκεφάλου, την εικόνα του διαρκώς μεταβαλλόμενου κόσμου (Churchland, 1994, p.6).

Η λειτουργική δραστηριότητα του εγκεφάλου όμως δεν συνίσταται στην απλή επανάληψη ή αναπαράσταση του αρχικού προτύπου ενεργοποίησης των εκπεμπόντων εικονοστοιχείων. Ένα νέο πρότυπο πρέπει να σχηματιστεί στο δέκτη μέσω της μεταβολής του εισερχόμενου προτύπου. Η διαδικασία εκ νέου αναπαράστασης του προτύπου ενεργοποίησης – την οποία ονομάζει *καταγραφή* ή *παρακολούθηση* - του προηγούμενου πληθυσμού νευρώνων αποτελεί για το συγγραφέα μια ενεργητική διαδικασία. Το αρχικό πρότυπο εικονοστοιχείων αρκετών νευρώνων του αμφιβληστροειδούς διαδιδόμενο «προς τα μέσα, από τον ένα εξειδικευμένο νευρωνικό πληθυσμό στον επόμενο, και από αυτόν στον επόμενο του, μετασχηματίζεται προοδευτικά σε κάθε στάδιο από τις παρεμβαλλόμενες διατάξεις συναπτικών συνδέσεων» (Churchland, 1994, p.11). Εκεί, λαμβάνει χώρα ο κύριος όγκος των υπολογισμών του εγκεφάλου και εκδηλώνονται ο χαρακτήρας και η ευφυΐα του ατόμου· «εδώ στηρίζεται τελικά η νοημοσύνη» (Churchland, 1994, p.11). Η παραπάνω διαδικασία φαίνεται στην παρακάτω εικόνα, όπου τα διαδοχικά κομμάτια των νευρώνων σχηματίζουν ένα νέο και διαφορετικό πρότυπο ενεργοποίησης.



Εικόνα 4: Οι διαδοχικοί μετασχηματισμοί των νευρωνικών προτύπων (*patterns*) ενεργοποίησης.

Ο Churchland διερευνά τα νευρωνικά δίκτυα ως μέσο για την κατανόηση των μηχανισμών που διέπουν τις γνωστικές διαδικασίες και τη συνείδηση. Αναγνωρίζει ότι τα νευρωνικά δίκτυα του εγκεφάλου παρουσιάζουν αξιοσημείωτη πολυπλοκότητα και δυνατότητες παράλληλης επεξεργασίας, γεγονός που τα καθιστά ένα συναρπαστικό μοντέλο για τη μελέτη του νου. Χρησιμοποιώντας τέτοια μοντέλα, οι ερευνητές μπορούν να προσομοιώσουν και να διερευνήσουν διάφορα γνωστικά φαινόμενα, όπως η αναγνώριση προτύπων, η γλωσσική επεξεργασία και η επίλυση προβλημάτων. Τα μοντέλα αυτά προσφέρουν έναν τρόπο για τη διερεύνηση της σχέσης μεταξύ της νευρωνικής δραστηριότητας και των γνωστικών λειτουργιών, παρέχοντας πληροφορίες για τους μηχανισμούς που οδηγούν σε συνειδητές εμπειρίες. Έμφαση δίνεται, ακόμα, στο ρόλο της κατανεμημένης επεξεργασίας και στις αναδυόμενες ιδιότητες που προκύπτουν από την αλληλεπίδραση πολλαπλών μονάδων επεξεργασίας. Η προοπτική αυτή ευθυγραμμίζεται με μια προσέγγιση από κάτω προς τα πάνω (bottom-up) για την κατανόηση του νου, όπου τα σύνθετα γνωστικά φαινόμενα προκύπτουν από τις αλληλεπιδράσεις απλούστερων υπολογιστικών στοιχείων.

4.3.2 Τα θεμελιώδη χαρακτηριστικά της συνείδησης

Ο Churchland φέρνει μια σειρά παραδειγμάτων γνωσιακών φαινομένων τα οποία μπορεί να πραγματοποιηθούν με επιτυχία με κάποιο αμιγώς φυσικό ή ηλεκτρονικό δίκτυο. Παρ'ολ'αυτά ο σκεπτικισμός σχετικά με την προοπτική της νευροφυσιολογικής αναγωγής παραμένει και εστιάζεται στο φαινόμενο της συνείδησης. Μπορεί ένα τέτοιο δίκτυο, παρά τις όποιες ικανότητές του, να λογίζεται ως ενσυνείδητο; Ο Churchland κρατάει 'μικρό καλάθι' για τις επιτυχίες της γνωσιακής νευροεπιστήμης μέχρις ότου οι διατυπώσεις της συμπεριλάβουν το νευραλγικό και ευαίσθητο κομμάτι της συνείδησης, αναλύοντας το πιο αινιγματικό μυστήριο με αμιγώς φυσικούς όρους.

Χωρίς υπεκφυγές αλλά κρατώντας αποφασιστική στάση, ο συγγραφέας επιλέγει να εξερευνήσει το ζήτημα προσδιορίζοντας ευθύς εξ'αρχής μερικές από τις χαρακτηριστικότερες ιδιότητες της συνείδησης. Θεωρεί λανθασμένο πρώτο βήμα την προσπάθεια έγκυρου ορισμού της συνείδησης, προκρίνοντας ως σημαντικότερη την αναζήτηση και αποσαφήνιση του τι είναι αυτό που πρέπει να αναδιατυπώσει η νευροεπιστήμη. Με δικά του λόγια, «οι ορισμοί διατυπώνονται καλύτερα, αφού προηγουμένως έχουμε καταλήξει σε επαρκή κατανόηση του τι χρειάζεται να οριστεί» (Churchland, 1994, p.213). Μας καλεί έτσι να εξετάσουμε τις ακόλουθες χαρακτηριστικές διαστάσεις της συνείδησης, υποστηρίζοντας ότι ένα κατάλληλα διαμορφωμένο δίκτυο επανατροφοδότησης θα επιδείξει γνωσιακές συμπεριφορές που αποτελούν λειτουργικά ανάλογα των παρακάτω διαστάσεων.

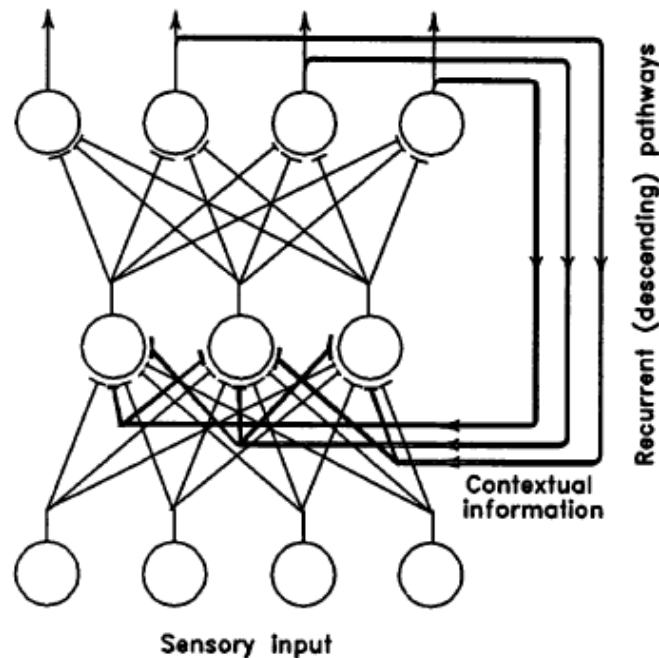
1. Η συνείδηση ενέχει τη βραχύχρονη μνήμη.

Για την ύπαρξη βραχύχρονης μνήμης απαιτείται τουλάχιστον κάποια γνωσιακή κατανόηση των γεγονότων που προηγούνται της παρούσας στιγμής. Η συνείδηση δημιουργεί την αίσθηση του τρόπου με τον οποίο οι εμπειρίες και το σώμα μας υπάρχουν

στο χρόνο, δηλαδή την ακολουθία των γεγονότων που ξετυλίγεται και συνιστά τον εκτεταμένο στο χρόνο κόσμο.

Η λειτουργία των οδών επανατροφοδότησης είναι η συνεχής μεταφορά, πίσω στη δεύτερη στιβάδα (hidden layer), επεξεργασμένων πληροφοριών που σχετίζονται με τις προηγούμενες καταστάσεις της ίδιας στιβάδας. Παρατηρείται στο εν λόγω σύστημα, η παρουσία μιας στοιχειώδους μορφής βραχύχρονης μνήμης.

Αυτή η εκπαιδεύσιμη μορφή βραχύχρονης μνήμης είναι άμεσα εξαρτώμενη από τη φύση της πληροφορίας, έχει μεταβλητό χρόνο φθοράς, ενώ αναδύεται από τη δομή και τη δυναμική ενός δικτύου επανατροφοδότησης, ως φυσικό χαρακτηριστικό τέτοιων συστημάτων. Ο Churchland, αν και αφήνει ανοιχτό το ενδεχόμενο αυτή η διαδικασία να αποτελεί και το υπόβαθρο της δικής μας βραχύχρονης μνήμης, θεωρεί τη συγκεκριμένη διαδικασία μια βιώσιμη εναλλακτική εξήγηση (Churchland, 1994, pp.213,216).



Εικόνα 5: Δίκτυο επανατροφοδότησης

2. Η συνείδηση είναι ανεξάρτητη από τα αισθητηριακά ερεθίσματα.

Ένα δίκτυο επανατροφοδότησης δεν έχει ανάγκη αισθητηριακής εισόδου, τουλάχιστον όσον αφορά τη συνεχή λειτουργία του. Μπορούμε να κλείσουμε τα μάτια και τα αυτιά μας, επιχειρώντας να ελαχιστοποιήσουμε όλες τις μορφές των αισθητηριακών μας ερεθισμάτων· η συνείδηση δε θα σβήσει με αυτόν τον τρόπο. Ακόμα θα μπορούμε να φανταστούμε το μέλλον, να αναπολήσουμε το παρελθόν ή να λύσουμε ένα σύνθετο πρόβλημα στο κεφάλι μας, χωρίς την επιρροή αισθητηριακών ερεθισμάτων. Αν και η παρατεταμένη αισθητηριακή στέρηση θα είχε αναμφίβολα αρνητικά αποτελέσματα στην ποιότητα της συνειδητής αντίληψης, η ύπαρξη της συνείδησης για μικρό χρονικό διάστημα δε φαίνεται να εξαρτάται από την ύπαρξη ή μη ερεθισμάτων.

Η συνεχής δραστηριότητά του εξασφαλίζεται από τα κωδικά διανύσματα που φτάνουν στη δεύτερη στιβάδα διαμέσου των επανατροφοδοτικών οδών. Το τυπικό αποτέλεσμα μιας τέτοιας διαδικασίας είναι μια συνεχής ακολουθία διανυσμάτων ενεργοποίησης που ξετυλίγονται στη δεύτερη στιβάδα, «μια καλά ορισμένη τροχιά στο χώρο ενεργοποίησης του δικτύου» (Churchland, 1994, p.217). Δεδομένης λοιπόν, της απουσίας - για κάποιο διάστημα - περιφερειακής αισθητηριακής εισόδου, πρέπει να περιγραφούν οι διανυσματικές τροχιές ή «οι γνωσιακές περιπλανήσεις που παράγονται εσωτερικά, ως ονειροπολήσεις, φαντασίες ή παθητικοί διαλογισμοί». Η γνωσιακή δραστηριότητα μπορεί, συνεπώς, να είναι αυτοπαραγόμενη (Churchland, 1994, p.217).

3. Η συνείδηση χαρακτηρίζεται από κατευθυνόμενη προσοχή.

Ένα βασικό χαρακτηριστικό της συνείδησης είναι η δυνατότητά της να εστιάζει σε μια αισθητηριακή οδό σε βάρος κάποιας άλλης, ακόμα κι όταν η εξωτερική αισθητηριακή προοπτική του κόσμου παραμένει σταθερή. Ένας αθλητής του μπάσκετ κατευθύνει όλη την προσοχή του στο καλάθι λίγο πριν εκτελέσει το τρίποντο, αγνοώντας άλλες πληροφορίες του περιβάλλοντος. Ο παρουσιαστής των ειδήσεων πρέπει να έχει συνεχώς τα μάτια του στραμμένα στον υποβολέα καταστέλλοντας λοιπές αισθητηριακές ροές. Ο νους βρίσκεται σε μια κατάσταση που παραμένει σταθερή κατά τη διάρκεια της ροής των αισθητηριακών ερεθισμάτων, μια κατάσταση που αυξάνει τη δυνατότητα αναγνώρισης κάποιων τύπων πραγμάτων έναντι άλλων. Από τη μια πλευρά, γεγονότα και καταστάσεις – που θα αναγνώριζες υπό άλλες συνθήκες - ξεφεύγουν της προσοχής σου· από την άλλη, αυτή η λειτουργία της συνείδησης επαυξάνει τοπικά, στο συγκεκριμένο θέμα όπου επικεντρώνεται, τη γνωσιακή επίδοση του ατόμου (Churchland, 1994, p.214).

Στα νευρωνικά δίκτυα, μας πληροφορεί ο Churchland, όταν επιθυμούμε να αυξήσουμε τις πιθανότητες συγκεκριμένης αναγνώρισης, πρέπει να αυξήσουμε την πιθανότητα της ενεργοποίησης του κατάλληλου πρωτοτυπικού διανύσματος από την αισθητηριακή είσοδο. Οι επανατροφοδοτικές οδοί επηρεάζουν αυτές τις πιθανότητες, προενεργοποιώντας τη σχετική στιβάδα προς την κατεύθυνση κάποιου πρωτοτυπικού διανύσματος (π.χ. τη σωστή τοποθέτηση του σώματος στην περίπτωση του μπασκετμπολίστα και την ανάγνωση του κειμένου από τον παρουσιαστή). Το δίκτυο εκείνη τη στιγμή ευνοεί εκείνο το συγκεκριμένο διάνυσμα, ως το τρέχον αντικείμενο της προσοχής του (Churchland, 1994, p.218).

4. Η συνείδηση έχει την ικανότητα να παράγει εναλλακτικές ερμηνείες σύνθετων ή ασαφών δεδομένων.

Σε περιπτώσεις όπου ένα ερέθισμα είναι συγκεχυμένο και προβληματικό, η συνείδηση συνεχίζει να παράγει και να εξερευνά ανταγωνιστικές ερμηνείες του περιεχομένου του. Τα δίκτυα επανατροφοδότησης έχουν την ικανότητα να υιοθετήσουν διαφορετικές γνωσιακές ερμηνείες της ίδιας αντιληπτικής περιστασης, λόγω του επανατροφοδοτικού χειρισμού της ίδιας τους της γνωσιακής επεξεργασίας. «Αυτή η ικανότητα αποτελεί συμπλήρωμα της ικανότητας για κατευθυνόμενη προσοχή» (Churchland, 1994, p.218). Στη μια περίπτωση, μια καθορισμένη νοητική κατάσταση επιβάλλεται σε μια διαρκώς μεταβαλλόμενη κατάσταση, προσπαθώντας να συλλάβει

σημαντικά χαρακτηριστικά όταν αυτά εμφανιστούν, ενώ στην άλλη (πολλαπλή ερμηνεία), η ψυχική διάθεση του ατόμου συνεχώς μεταβάλλεται και προσαρμόζεται σχετικά με μια προβληματική κατάσταση. Και τα δυο αυτά φαινόμενα προκύπτουν ως φυσικές συνέπειες των δικτύων.

5. Η συνείδηση εξαφανίζεται στο βαθύ ύπνο.

Ο πιο κοινός τρόπος απώλειας της συνείδησης είναι η κατάσταση του βαθύ ύπνου. Πώς εξηγείται αυτή η απώλεια και η επανεμφάνιση της συνείδησης όταν ονειρευόμαστε, κατά τη διάρκεια του σταδίου REM (rapid eye movement);

Παρουσιάζοντας τα ευρήματα μιας επιστημονικής έρευνας του Rodolfo Llinas, ο Churchland περιγράφει τρεις διαφορετικές καταστάσεις της δραστηριότητας του εγκεφάλου. Με τη χρήση της μαγνητοεγκεφαλογραφίας (MEG), ο Llinas κατάφερε να καταγράψει τη συλλογική δραστηριότητα των δισεκατομμυρίων νευρώνων σε όλο τον εγκεφαλικό φλοιό (Churchland, 1994, p.219).

Κατά τη διάρκεια της λειτουργίας της συνείδησης ενός φυσιολογικού ανθρώπου σε εγρήγορση, η λειτουργία αυτή χαρακτηρίζεται από μια σταθερή, βασική συχνότητα 40 Hz, η οποία επικαλύπτεται από μεγάλες μη περιοδικές εναλλαγές στα επίπεδα της νευρικής δραστηριότητας (εικόνα 6, πρώτο διάγραμμα). Προκύπτει ότι κατά τη διάρκεια της φυσιολογικής λειτουργίας της συνείδησης υπό εγρήγορση, οι εξάρσεις δραστηριότητας συσχετίζονται στενά με τις αλλαγές στο αντιληπτικό περιβάλλον του υποκειμένου. Η γνωσιακή δραστηριότητα που ανιχνεύεται στο φλοιό είναι μια εκτυλισσόμενη αναπαράσταση, εν μέρει του εκτυλισσόμενου αντιληπτικού περιβάλλοντος του υποκειμένου (Churchland, 1994, p.220).

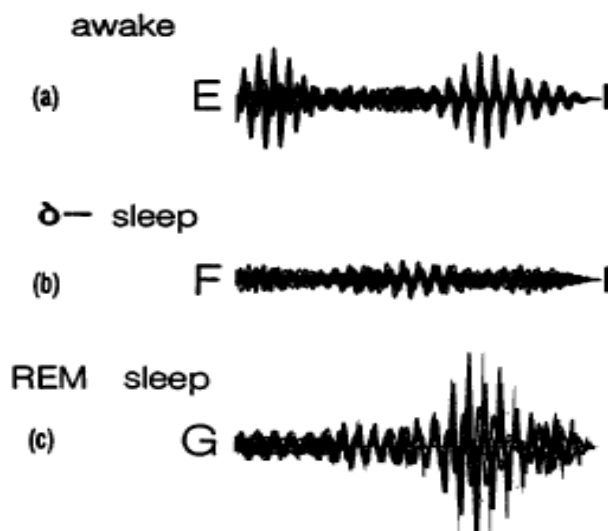
Στο δεύτερο διάγραμμα της εικόνας 6 απεικονίζεται η δραστηριότητα του εγκεφάλου κατά τη διάρκεια του βαθύ ύπνου, όπου η ταλάντωση των 40 Hz παραμένει αν και με αισθητά μικρότερο πλάτος. Οι εξάρσεις όμως φαίνεται να απουσιάζουν. Η ισχυρή αναπαραστατική δραστηριότητα της κατάστασης εγρήγορσης έχει εξαφανιστεί. Το μεγάλο υποσύστημα του εγκεφάλου βρίσκεται εκτός λειτουργίας.

6. Η συνείδηση επανεμφανίζεται όταν ονειρευόμαστε, σε μια βουβή ή κομματιασμένη, τουλάχιστον, μορφή.

Κατά τη διάρκεια των ονείρων, παρατηρούμε την επανεμφάνιση της ισχυρής αναπαραστατικής δραστηριότητας, με τις χαρακτηριστικές μη περιοδικές ταλαντώσεις να επικαλύπτουν τη συχνότητα των 40 Hz (εικόνα 6, τρίτο διάγραμμα). Η θεμελιώδης διαφορά με την κατάσταση εγρήγορσης έγκειται στο ότι κατά τη διάρκεια του σταδίου REM, η αναπαραστατική δραστηριότητα δε συσχετίζεται με αλλαγές στο περιβάλλον του υποκειμένου. Η αλλαγή της φωτεινότητας του δωματίου για παράδειγμα, ενώ στην προηγούμενη κατάσταση εγρήγορσης θα αποτυπωνόταν σαν έξαρση των ταλαντώσεων, στο παρόν στάδιο REM δεν καταγράφεται στη ροή της νευρωνικής δραστηριότητας αυτού που ονειρεύεται. Το σενάριο της αναπαραστατικής ιστορίας πλέον πλάθεται από εσωτερικούς παράγοντες και όχι από εξωτερικά ερεθίσματα.

Η φύση της ονειρικής δραστηριότητας μπορεί ακόμα να εξηγήσει τον κοινό και πρωτοτυπικό χαρακτήρα των ονείρων. Ο κύριος παράγοντας που καθορίζει τις περιπλανώμενες τροχιές του συστήματος στον χώρο των ενεργοποιήσεων είναι το προηγούμενο τοπίο των χρονικά δομημένων πρωτοτύπων που προϋπάρχουν, ελλείψει του ελέγχου που πραγματοποιούν τα αισθητηριακά ερεθίσματα στο επανατροφοδοτικό σύστημα. Καθοριστικό παράγοντα συνιστά επίσης, η γνωσιακή και συναισθηματική κατάσταση του υποκειμένου πριν κοιμηθεί και ο χαμηλού επιπέδου ενεργοποιητικός θόρυβος.

Ο Churchland τονίζει επίσης τη σημασία του ενδοπετάλιου πυρήνα στην εμφάνιση της ενσυνείδητης γνωσιακής δραστηριότητας. Ενδεχόμενη βλάβη στον ενδοπετάλιο δύναται να παράξει βαθύ και μη αναστρέψιμο κώμα, οδηγώντας σε πλήρη εξαφάνιση της συνείδησης. «Το επανατροφοδοτικό σύστημα δεν μπορεί να ασχοληθεί με τις πολύπλοκες δραστηριότητές του, εάν το στενό εκείνο σημείο του συστήματος τεθεί εκτός λειτουργίας» (Churchland, 1994, p.220).



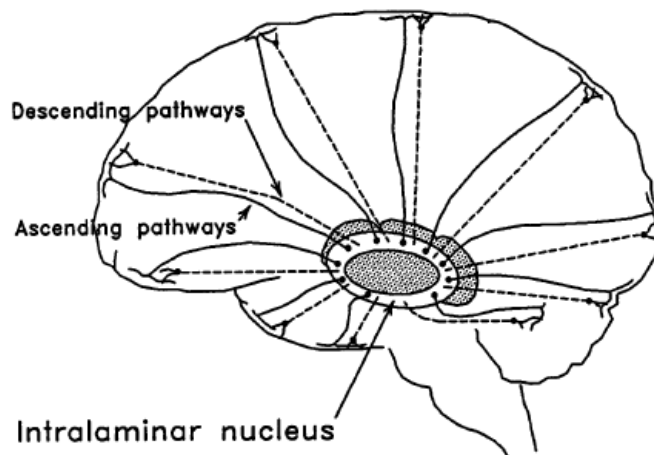
Εικόνα 6: Τρεις διαφορετικές καταστάσεις της δραστηριότητας του εγκεφάλου: a) Η φλοιώδης δραστηριότητα κατά τη διάρκεια της εγρήγορσης. (b) Η φλοιώδης δραστηριότητα κατά τη διάρκεια του βαθύ ύπνου. (c) Η φλοιώδης δραστηριότητα κατά τη διάρκεια του REM ύπνου.

7. Η συνείδηση συνενώνει τα περιεχόμενα διάφορων βασικών αισθητηριακών τύπων μέσα σε μια μοναδική και ενοποιημένη εμπειρία.

Στον ενδοπετάλιο πυρήνα βρίσκεται ένα ευρέως εκτεταμένο επανατροφοδοτικό σύστημα με ένα στενό σημείο κυκλοφορίας των πληροφοριών (εικόνα 7). Τα διανύσματα ενεργοποίησης του ενδοπετάλιου και η ακτινική νευροαξονική δραστηριότητα αναπαριστούν τις πληροφορίες από τις αισθητηριακές φλοιώδεις περιοχές που εισέρχονται στο επανατροφοδοτικό σύστημα. «Οι αναπαραστάσεις σε αυτό το επανατροφοδοτικό σύστημα πρέπει, συνεπώς, να είναι πολυτροπικές» (Churchland, 1994, p.222).

Ο βασικός ισχυρισμός του συγγραφέα είναι ότι μια γνωσιακή αναπαράσταση αποτελεί στοιχείο της τρέχουσας συνειδησιακής κατάστασης, υπό την απαραίτητη

προϋπόθεση ότι είναι ένα διάνυσμα ενεργοποίησης ή μια ακολουθία τέτοιων διανυσμάτων μέσα στο ευρύ επανατροφοδοτικό σύστημα της εικόνας 7 .



Εικόνα 7: Οι αξονικές προβολές που συνδέουν όλες τις περιοχές του εγκεφαλικού φλοιού με τον ενδοπετάλιο πυρήνα του θαλάμου. Με διακεκομμένες γραμμές σημειώνονται οι επιστρέφουσες νευρικές οδοί.

Ένα πρόβλημα που αναγνωρίζει ο Churchland ότι προκύπτει από την παραπάνω ανάλυσή του, είναι ότι το μεγάλο επανατροφοδοτικό κύκλωμα νευρικών οδών από τον φλοιό στον ενδοπετάλιο πυρήνα μπορεί να είναι πολύ αραιό για να μεταφέρει το πλούσιο πληροφοριακό φορτίο που φαίνεται να απαιτεί η συνείδηση. Ενδέχεται η λειτουργία των οδών να περιορίζεται στη μέτρηση του χρόνου. Ανεξάρτητα όμως από τα προβλήματα που μπορεί να προκύψουν στο μέλλον, η εξηγητική ισχύς των επιχειρημάτων του Churchland βρίσκεται στις προτάσεις του για τις δυναμικές ιδιότητες των επανατροφοδοτικών δικτύων (Churchland, 1994, p.223).

«Μπορούμε τώρα να δούμε», δηλώνει ο Churchland, «ότι δεν υπάρχει τίποτα αποκλειστικά αντικειμενικό στα φυσικά φαινόμενα, στο βαθμό που μπορεί να γίνουν εξίσου γνωστά με υποκειμενικά μέσα, συγκεκριμένα διαμέσου της δραστηριότητας των αυτοσυνδεδεμένων επιστημικών οδών». Οι φυσικές καταστάσεις του εγκεφάλου δεν είναι περισσότερο αντικειμενικές από ότι η φυσική ύλη του σώματος είναι εγγενώς νεκρή. Η λειτουργία του οργανωμένου φυσικού συστήματος καθορίζει και τα δυο (Churchland, 1994, p.225).

Παρόμοια, οι νοητικές μας καταστάσεις δε διαθέτουν κάτι αποκλειστικά υποκειμενικό, καθώς μπορούν να γίνουν γνωστές μέσω διαφόρων πληροφοριακών οδών· η νοητική μας κατάσταση εκδηλώνεται μέσα στο λόγο, στις εκφράσεις του προσώπου και στη συνολική φυσική συμπεριφορά μας. Η φιλοσοφική θέση του Churchland διατυπώνεται εύγλωττα στο ότι «απλούστατα, δεν υπάρχει σύγκρουση μεταξύ του να είναι κάτι ταυτόχρονα και αντικειμενικό και υποκειμενικό», καθώς «η ίδια κατάσταση μπορεί να είναι και τα δυο» (Churchland, 1994, p.220).

4.4 Η γλωσσική λειτουργία στον Churchland

Κατά τη διαδικασία σύγκρισης της ανθρώπινης νοημοσύνης με τη νοημοσύνη άλλων ειδών, αυτό που πρέπει να λαμβάνεται υπόψη είναι τα δυο διανύσματα των γνωσιακών ικανοτήτων που χαρακτηρίζουν κάθε είδος. Χρειάζεται δηλαδή να συγκριθούν δύο σύνθετα πρότυπα ή προφίλ τα οποία αλληλεπικαλύπτονται σε κάποιο βαθμό.

Μια αναμφισβήτητη γνωσιακή εξειδίκευση του ανθρώπου είναι η γλώσσα. Υπάρχουν στο υπόλοιπο ζωικό βασίλειο στοιχειώδεις ενδείξεις του ίδιου είδους συστηματικής ικανότητας που αποτελεί τη βάση της ανθρώπινης γλώσσας; Παρακάτω εξετάζονται διαφορετικές οπτικές σε σχέση με αυτή τη συνειδητή λειτουργία.

4.4.1 Κριτική στην έννοια του γλωσσικού οργάνου

Σύμφωνα με τον διάσημο γλωσσολόγο Noam Chomsky, ο ανθρώπινος εγκέφαλος είναι προικισμένος με ένα ειδικό γλωσσικό όργανο, προνόμιο αποκλειστικά των ανθρώπων. Ουσιαστικά, πρόκειται για ένα νευρωνικό υποσύστημα το οποίο διαθέτει την έμφυτη ικανότητα χειρισμού των διαφόρων γραμματικών κανόνων σύμφωνα με τους οποίους λειτουργούν όλα τα ανθρώπινα γλωσσικά συστήματα. Αυτό συνεπάγεται ότι κανένα άλλο ζώο δε θα μπορούσε να αναπτύξει τις δικές μας γλωσσικές δεξιότητες, τη συστηματική ικανότητα που αποτελεί τη βάση της ανθρώπινης γλώσσας.

Οι θεωρίες του Dennett και του Chomsky οδηγούν στην υιοθέτηση της άποψης ότι η συνείδηση των ανθρώπων διαφέρει εκ βάθρων από αυτή των υπόλοιπων ζώων, αφής στιγμής που στερούνται τη γλωσσική ικανότητα. Ο Churchland στην προσπάθειά του να διαψεύσει τον ισχυρισμό του Chomsky - περί ύπαρξης κάποιου ειδικού γλωσσικού οργάνου - παραθέτει μια σειρά ανταγωνιστικών εμπειρικών δεδομένων.

Μέσα από τη μελέτη των κακώσεων, προκύπτει -σε αντίθεση με τις αρχικές ενδείξεις- ότι οι γλωσσικές δεξιότητες δεν περιορίζονται σε μια εξειδικευμένη περιοχή του εγκεφάλου, αλλά κατανέμονται σε ευρύτερες περιοχές: διαφορετική περιοχή για την πρόσβαση στα ρήματα, διαφορετική για τα κύρια ονόματα, για τα ουσιαστικά, διαφορετική για την κατανόηση της γραμματικής και διαφορετική για την παραγωγή της. Η περιοχή Broca στην οποία αρχικά αποδιδόταν ο ρόλος του 'γραμματικού κουτιού', τελικά φαίνεται να είναι υπεύθυνη κυρίως για τη γνώση των ρημάτων. «Εάν υπάρχει κάποιο ιδιαίτερο γλωσσικό όργανο, είναι διασκορπισμένο στο μεγαλύτερο μέρος της επιφάνειας του εγκεφάλου και βρίσκεται σε περιοχές για τις οποίες πολλά από τα άλλα ζώα έχουν σαφή ανατομικά ομόλογα», σχολιάζει ο Churchland (Churchland, 1994, p.260).

Μια ακόμα προβληματική τίθεται με την ύπαρξη δίγλωσσων και πολύγλωσσων ανθρώπων. Ακόμα κι αν δεχτούμε ότι το ίδιο όργανο μπορεί να χειριστεί αρκετές διαφορετικές γλώσσες, μέσω της κατοχής μιας μοναδικής σύνθετης γλώσσας με διαφορετικές μορφές λόγου, τα προβλήματα επανέρχονται εξετάζοντας ορισμένες κλινικές περιπτώσεις. Χαρακτηριστικό παράδειγμα αποτελεί η περίπτωση ενός δίγλωσσου Ελληνοαμερικάνου που έχασε την ικανότητα χειρισμού της ελληνικής γλώσσας, κατά τη διάρκεια τεχνητής καταστολής περιοχών του αριστερού ημισφαιρίου. Η γνώση της αγγλικής, ωστόσο, παρέμεινε ανέπαφη, ενώ το αντίθετο συνέβη με την καταστολή γειτονικών περιοχών. Αυτές οι μελέτες υποδηλώνουν ότι οι δίγλωσσοι πρέπει να διαθέτουν

τουλάχιστον δυο γλωσσικά όργανα σε ξεχωριστό χώρο και με ξεχωριστή λειτουργία, αμφότερα στην ίδια πλευρά του εγκεφάλου (Churchland, 1994, p.261).

Σύνηθες είναι ακόμα, στην περίπτωση που μια εγκεφαλική βλάβη εμποδίζει τις δομές του αριστερού ημισφαιρίου να εκτελούν τις φυσιολογικές γλωσσικές λειτουργίες τους, οι αντίστοιχες περιοχές του δεξιού ημισφαιρίου να αναλάβουν αυτό το ρόλο. Σε ένα φυσιολογικό παιδί που υποφέρει από μια τέτοια βλάβη, η γλωσσική ικανότητα αναπτύσσεται εντός των δομών του δεξιού ημισφαιρίου του εγκεφάλου έναντι του αριστερού. Υπάρχει συνεπώς, κατά μια έννοια, γλωσσικό όργανο και στη δεξιά πλευρά του εγκεφάλου. Ωστόσο, η υπόθεση ότι αυτό το όργανο είναι υπολογιστικά αφιερωμένο στην εφαρμογή και το χειρισμό των αφηρημένων γραμματικών κανόνων καθίσταται προβληματική, καθώς η συνήθης λειτουργία των δομών του δεξιού ημισφαιρίου σχετίζεται με διαφορετικά γνωσιακά καθήκοντα (Churchland, 1994, p.261).

Ο Churchland καταλήγει ότι «είναι πολύ πιθανότερο, σύμφωνα με την άποψη αυτήν, να έμαθαν οι άνθρωποι σε κάποιο σημείο να κάνουν πλήρη ή καινούργια χρήση μιας υπολογιστικής ικανότητας που όλα τα πρωτεύοντα θηλαστικά κατέχουν, τουλάχιστον σε κάποιο βαθμό» (Churchland, 1994, p.261).

4.4.2 Κριτική στη γλωσσοκεντρική θεωρία του Dennett

Ο Dennett συντάσσεται με την πλευρά των επιστημόνων - ανάμεσά τους και ο Churchland - που θεωρεί ότι η αρχιτεκτονική του εγκεφάλου έχει τη μορφή ενός μαζικά παράλληλου υπολογιστικού συστήματος. Παρ'ότι δεν διαθέτει τη δομή ενός κλασικού, σειραϊκού υπολογιστή διακριτών καταστάσεων, ο ανθρώπινος εγκέφαλος μπορεί να προσομοιώσει τη συμπεριφορά μιας τέτοιας κλασικής μηχανής, όπως π.χ. παραγωγή και κατανόηση σύνθετων σειρών γλωσσικών συμβόλων, πραγματοποίηση αξιόπιστων παραγωγικών και αριθμητικών πράξεων. Ο Dennett θεωρεί ότι κατά τη διάρκεια της εκτέλεσης αυτών των διεργασιών, η υποκείμενη παράλληλη αρχιτεκτονική των νευρώνων συνιστά μια 'εικονική' υπολογιστική μηχανή με δραστηριότητες κλασικού, διακριτών καταστάσεων, καθοδηγούμενου από κανόνες, σειραϊκού τύπου (Churchland, 1994, p.264).

Αυτή η μετατροπή του υπολογιστή σε εξειδικευμένη εικονική μηχανή σχετίζεται με το πρόγραμμα που εκτελείται εκείνη τη στιγμή. Η εγκατάσταση, για παράδειγμα, του προγράμματος Word μετατρέπει τον υπολογιστή σε μηχανή επεξεργασίας κειμένου, η εγκατάσταση του Ansys σε μηχανή προσομοίωσης μοντέλων μηχανικής κλπ. Συνεπώς, με τον κατάλληλο προγραμματισμό του, μπορούμε να προσομοιώσουμε ακόμα και τη συμπεριφορά ενός μαζικά παράλληλου νευρωνικού δικτύου.

Με παρόμοιο τρόπο, ένα πραγματικό, μαζικά παράλληλο νευρωνικό δίκτυο μπορεί – κατά τον Dennett – να προσομοιώσει της υπολογιστικές δραστηριότητες μιας σειραϊκής μηχανής, εφόσον ρυθμιστούν κατάλληλα τα συναπτικά του βάρη. Αυτή η μετατροπή λαμβάνει χώρα και στη διαδικασία της γλωσσικής μάθησης από τον ανθρώπινο εγκέφαλο, ο οποίος αναπτύσσει την ικανότητα «να αναπαριστά και να επεξεργάζεται πληροφορίες με μια δομημένη σειρά αναπαραστάσεων που καθοδηγούνται από κανόνες και εκτυλίσσονται στο χρόνο». Για τον Dennett η ροή της ανθρώπινης συνείδησης συνίσταται ακριβώς σε αυτή

την εκτυλισσόμενη ακολουθία αναπαραστάσεων, στην ευρεία ροή γλωσσικής δραστηριότητας. Από τη στιγμή που αυτές οι, κυρίως γλωσσικές, δραστηριότητες (σειραϊκές δραστηριότητες διακριτών καταστάσεων, κλπ) εντάσσονται στην ανθρώπινη σφαίρα εκμάθησης ικανοτήτων, η ειδικού τύπου συνείδηση που απολαμβάνουμε δεν μπορεί να υπάρξει και στα υπόλοιπα ζώα (Churchland, 1994, p.264).

Ο Churchland αντιτίθεται στο πρωτότυπο της γλωσσικής δραστηριότητας που έχει κυριαρχήσει στις θεωρητικές προσπάθειες εξήγησης της ανθρώπινης ενσυνείδητης εμπειρίας. Αντ'αυτού θεωρεί κρισιμότερη τη συμβολή της θεωρίας της κατανεμημένης διανυσματικής επεξεργασίας σε ένα μαζικά παράλληλο νευρωνικό δίκτυο επανατροφοδότησης, ως την κύρια μορφή υπολογιστικής δραστηριότητας σε όλους τους βιολογικούς εγκεφάλους. Σε αυτό το πλαίσιο κατηγορεί τον Dennett ότι «επιχειρεί να επαναφέρει αυτό το αποτυχημένο πρωτότυπο στο φως της επικαιρότητας, το μετατρέπει σε μοντέλο της ανθρώπινης συνείδησης, επιδοκιμάζει επιπόλαια την παράλληλα κατανεμημένη επεξεργασία, γιατί μπορεί να προσομοιώσει μια 'εικονική περίπτωση' του παλαιού γλωσσικού πρωτοτύπου και αντιμετωπίζει την αδυναμία της θεωρίας του να εξηγήσει τη συνείδηση των μη γλωσσικών πλασμάτων αρνούμενος ότι αυτά έχουν κάτι που να μοιάζει έστω με την ανθρώπινη συνείδηση» (Churchland, 1994, p.266).

Στην προσομοίωση ενός παράλληλου νευρωνικού δικτύου σε μια κλασική σειραϊκή μηχανή, η μηχανή δεν μας δίνει κάτι πέραν της αφηρημένης συμπεριφοράς εισόδου-εξόδου του προσομοιωμένου δικτύου. Ο κλασικός υπολογιστής δεν μπορεί να πραγματοποιήσει καμιά γνησίως παράλληλη κωδικοποίηση ή επεξεργασία, εφόσον το εσωτερικό του υπακούει στην άκαμπτη σειραϊκή μορφή των διακριτών καταστάσεων. Γενικεύοντας το συμπέρασμα αυτού του παραδείγματος, ο Churchland δηλώνει ότι «όταν λέμε ότι κάποια κλασική μηχανή **M** έχει προγραμματιστεί έτσι ώστε να λειτουργεί ως κάποια εικονική μηχανή με εξειδικευμένη λειτουργία **V**, τίποτα δε συνάγεται από αυτό σχετικά με τις εσωτερικές υπολογιστικές διαδικασίες της **M** και σχετικά με τον τρόπο με τον οποίο αυτές μπορεί να μιμούνται τις εσωτερικές διαδικασίες της μηχανής-στόχου που προσομοιώνεται, εκτός από το ότι καταφέρνουν με κάποιον τρόπο να παράγουν τις ίδιες συμπεριφορές εισόδου-εξόδου» (Churchland, 1994, pp.266-267). Αντίστοιχα, στην αντίθετη κατεύθυνση, όταν μια παράλληλη μηχανή προσομοιώνει επιτυχώς τη συμπεριφορά μιας σειραϊκής, αυτό απλά σημαίνει ότι έχει επιτευχθεί η σχετική συμπεριφορά εισόδου εξόδου, χωρίς να μπορεί να γίνει λόγος για τις εκτυλισσόμενες εσωτερικές διεργασίες της.

Σύμφωνα με τον Churchland, για να εξηγηθεί το φαινόμενο της συνείδησης, απαιτείται η αυθεντικά παράλληλη λειτουργία και όχι απλά μια προσομοίωση εισόδου-εξόδου. Η ροή της συνείδησης είναι ουσιαστικά απύσχα από τη στιγμή που το παράλληλο σύστημα του εγκεφάλου αποτυγχάνει να εκτελέσει αυθεντικές σειραϊκές υπολογιστικές δραστηριότητες. Ενώ θα συμφωνήσει με τον Dennett στο ότι η ανθρώπινη συνείδηση ενέχει την ομαλή ροή των γνωσιακών δομών στο χρόνο, θεωρεί ότι αυτός σφάλει όταν πιστεύει ότι «οι κλασικές γλωσσικές υπολογιστικές διεργασίες προσφέρουν την καλύτερη ελπίδα κατανόησης αυτής της ροής στο χρόνο» (Churchland, 1994, p.267).

Δεν υπάρχει κάποια ανάγκη να εισάγουμε τις κλασικές διεργασίες ή τη γλωσσική επεξεργασία για να εξηγήσουμε την ομαλή χρονική ροή της συνείδησης, κατά τον Churchland. Η δυναμική συμπεριφορά των πραγματικών δικτύων επανατροφοδότησης με την ικανότητα τους για παραγωγή σύνθετων αναπαραστάσεων με μια διαρκώς εκτυλισσόμενη χρονική διάσταση, μας παρέχει ευρύτερες και αποτελεσματικότερες λύσεις. Η χρονικά δομημένη δραστηριότητα είναι η φυσική ιδιότητα κάθε δικτύου επανατροφοδότησης, ανεξάρτητα από τις γλωσσικές δεξιότητες που έχει αποκτήσει. Επιπλέον, αυτή η ικανότητα σύνθεσης δομών στο χρόνο προσφέρει μια ενοποιημένη εξήγηση της συνείδησης σε όλα τα ανώτερα ζώα, ανεξάρτητα από το αν χρησιμοποιούν γλώσσα ή όχι. Η θεωρία του Dennett πέρα από ανεπαρκής στην εξήγηση της ανθρώπινης συνειδητής λειτουργίας, αποδεικνύεται και άδικη προς το υπόλοιπο ζωικό βασίλειο προσπερνώντας το γεγονός ότι τα περισσότερα ζώα έχουν φλοιό με πολλές στιβάδες, εσωτερικά συνδεδεμένες βρεγματικές αναπαραστάσεις και εκτεταμένες επανατροφοδοτικές συνδέσεις μεταξύ θαλάμου και φλοιού, όπως οι άνθρωποι.

Συνολικά, η θεωρία του Dennett κρίνεται από τον Churchland ως ανεπαρκής για τις επεξηγηματικές της δυνατότητες. Περιορίζει το πολύπλευρο φαινόμενο της συνείδησης στη στενά γλωσσική του μορφή. Με αυτό τον τρόπο, μια σειρά οπτικών, μουσικών, απτικών, κινητικών, κοινωνικών ακολουθιών μένουν εκτός της περιγραφής των συνειδητών λειτουργιών.

4.4.3 Η σημασία της γλώσσας στη σκέψη και στη συνείδηση

Ενώ ο κοινωνικός θεσμός της γλώσσας δε σχετίζεται με τη γέννηση του φαινομένου της συνείδησης, το περιεχόμενο της συνείδησης διαμορφώνεται σε μεγάλο βαθμό από το κοινωνικό περιβάλλον στο οποίο αναπτύσσεται η συνείδηση. Το πολιτισμικό πλαίσιο, εντός του οποίου μεγαλώνει και κοινωνικοποιείται ένα παιδί, επιβάλλει την εκμάθηση των αιτιωδών και κοινωνικών κατηγοριών, δεδομένων για το υπόλοιπο κοινωνικό σύνολο. Αυτές οι κατηγορίες αντικατοπτρίζονται στη γλώσσα και εσωτερικεύονται κατά τη διαδικασία εκμάθησής της (Churchland, 1994, p.269).

Το παιδί, με αυτόν τον τρόπο, κατακτά τη συσσωρευμένη πείρα του παρελθόντος, καθώς οι περισσότερες - αποκτημένες σε προηγούμενες γενιές - γνώσεις για τον κόσμο αντανakλώνται στο λεξιλόγιό του. Οι σχετικές έννοιες και τα πρωτότυπα που πρέπει ακόμα να μάθει θα προκύψουν από τις συνεχείς αλληλεπιδράσεις του με τον κόσμο. Το λεξιλόγιο αποτελεί ένα περίγραμμα εντός του οποίου διαμορφώνεται η ανάπτυξη του εγκεφάλου του παιδιού (Churchland, 1994, p.269).

«Η γλώσσα, λοιπόν, αποτελεί μια μορφή εξωσωματικής μνήμης, ένα μέσο αποθήκευσης πληροφοριών, που υπάρχει έξω από κάθε εγκέφαλο και επιζεί μετά το θάνατο του εγκεφάλου», σημειώνει ο Churchland. Η αποκτημένη πληροφορία, μέσω των κατάλληλων μετασχηματισμών, μεταδίδεται από γενιά σε γενιά, ενώ η γραπτή αποτύπωσή της εντείνει αυτή τη διαδικασία. Η εύρεση λύσεων στα γνωσιακά προβλήματα, πλέον, δεν περιορίζεται στη μνήμη ενός ατόμου, καθώς η γλώσσα καθιστά δυνατή τη συλλογικότητα του ανθρώπινου γινώσκειν συνεισφέροντας στην υπέρβαση των ατομικών γνωσιακών αδυναμιών (Churchland, 1994, p.270).

Η συλλογικοποίηση της γνωσιακής δραστηριότητας και η έκταση στο χρόνο, αποτελούν δυο διαστάσεις της γλωσσικής λειτουργίας που δίνουν πλεονέκτημα στα είδη που χρησιμοποιούν κάποιου είδους γλώσσα. Τα περιεχόμενα της ανθρώπινης συνείδησης υπερβαίνουν συχνά την επινοητικότητα και τη φαντασία κάθε ζώου. Παρ'όλ'αυτά, «η συνείδηση αυτή καθαυτή δε χρειάζεται να είναι τίποτα διαφορετικό στον άνθρωπο από ότι στα άλλα ζώα». Αν και ένας βαθύς μετασχηματισμός των περιεχομένων της ανθρώπινης συνείδησης συντελέστηκε με τη γλώσσα και συνεχίζει να συντελείται, η φύση του ίδιου του φαινομένου της συνείδησης φαίνεται να είναι κοινό κτήμα μεταξύ των περισσότερων ειδών του ζωικού βασιλείου (Churchland, 1994, p.271).

4.5 Η εννοιολογική διεύρυνση ⁷ και η αλλαγή παραδείγματος

Αλλά πέρα από τη συνάφεια μας με τα υπόλοιπα ζώα, ποια επιτεύγματα και τεχνικές μας κάνουν ξεχωριστούς; Ο Churchland βρίσκει την απάντηση στα εννοιολογικά και πρακτικά πλαίσια των θεωρητικών και πειραματικών επιστημών που αναπτύχθηκαν σε βάθος χρόνων. Τα οικοδομήματα της σύγχρονης φυσικής, της χημείας, της βιολογίας, της ιατρικής, μας επιτρέπουν να γνωρίζουμε τη χημική σύσταση των αστεριών, να κατασκευάζουμε σύνθετα υλικά αλλά και να θεραπεύουμε ασθένειες άλλοτε ανίατες. Αναρωτώμενος για το πώς ο άνθρωπος κατάφερε να φτάσει σε αυτό το βάθος ουσιαστικής κατανόησης των φυσικών φαινομένων, ο Churchland εξετάζει το έργο ενός άλλου επιστήμονα και φιλοσόφου της επιστήμης, του Thomas Kuhn.

4.5.1 Η κατανόηση της θεωρίας μέσω «παραδειγμάτων»

Η «Δομή των Επιστημονικών Επαναστάσεων» (1962), το δημοφιλέστερο έργο του Kuhn, ταρακούνησε ολόκληρη τη φιλοσοφική κοινότητα. Ο Churchland αποδίδει τη μεγάλη αντιπαράθεση που ακολούθησε σε δυο λόγους. Κατ'αρχάς, μια βασική θέση του βιβλίου είναι ότι οι επιστημονικές επαναστάσεις του παρελθόντος δεν ήρθαν ως συνέπεια καθαρά λογικών και πειραματικών παραγόντων, οι οποίοι «έπαιξαν ορθολογικά το ρόλο τους σύμφωνα με κάποια καλά ορισμένη μεθοδολογία». Περισσότερο οι επαναστάσεις ήταν προϊόν μιας σειράς κοινωνικών, ψυχολογικών, μεταφυσικών, τεχνολογικών, αισθητικών και προσωπικών παραγόντων. Ο βαθμός στον οποίο η λογική διαδραμάτισε κάποιο ρόλο στην έκβαση αυτών των επαναστάσεων έχει, σύμφωνα με τον Kuhn, υπερτονιστεί από την ιστορία της επιστήμης που ακολούθησε. Ο Churchland χαρακτηριστικά διαπιστώνει ότι «οι ηρωικές και καθαρά λογικές ιστορίες γράφονταν πάντα, φυσικά, από τους νικητές της αρχικής σύγκρουσης ή από τους διανοητικούς απόγονούς τους». Λογικές έννοιες όπως επαγωγές, απορρίψεις, επικυρώσεις δεν επιτρέπουν την επαρκή κατανόηση των πραγματικά καθοριστικών παραγόντων (Churchland, 1994, p.272).

Ο δεύτερος λόγος της φιλοσοφικής διαμάχης ήταν η καλά τεκμηριωμένη θέση, ότι η μονάδα της επιστημονικής κατανόησης δεν είναι η πρόταση ή το σύνολο προτάσεων, αλλά

⁷ Περισσότερα για την εννοιολογική διεύρυνση στο υποκεφάλαιο 4.5.3

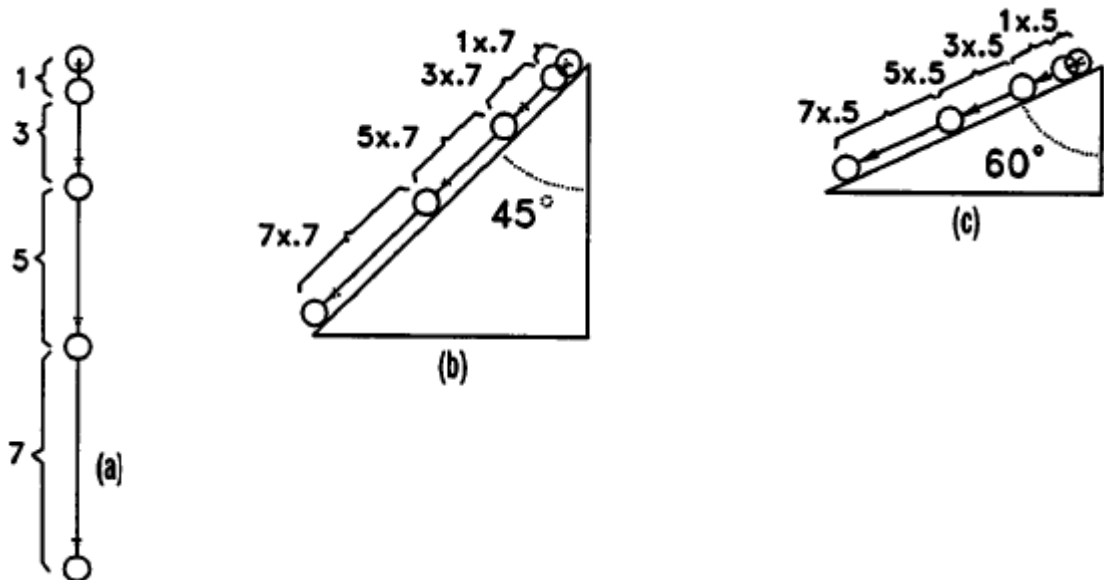
το ‘παράδειγμα’ (paradigm) ή ομάδα παραδειγμάτων. Ο όρος ‘παράδειγμα’ στο έργο του Kuhn έχει την έννοια του συγκεκριμένου παραδείγματος για τον τρόπο κατανόησης κάποιου πράγματος. Από αυτήν την πρωτοτυπική εξήγηση μπορούμε να παράξουμε άλλες σχετικές εξηγήσεις στο συγκεκριμένο πεδίο, ως παραλλαγές στο βασικό θέμα. Στην εκμάθηση μιας επιστημονικής θεωρίας, ο μαθητής ξεκινά από τα πρωτοτυπικά χαρακτηριστικά του κεντρικού παραδείγματος για να επεκτείνει αυτήν την κατανόηση και στα υπόλοιπα παραδείγματα «που απλώνονται ακτινωτά από το κεντρικό παράδειγμα που ήδη κατέχει» (Churchland, 1994, p.272).

Για να εντυπωθεί ευκρινέστερα στον αναγνώστη η έννοια του «παραδείγματος», ο Churchland φέρνει μια σειρά παραδειγμάτων από τη μηχανική. Στο σχήμα (a) απεικονίζεται ένα σώμα που εκτελεί ελεύθερη πτώση, κινείται δηλαδή σε ευθεία γραμμή προς τα κάτω με αυξανόμενη ταχύτητα. Οι διαδοχικές κατακόρυφες αποστάσεις που διανύει, σε ίσα χρονικά διαστήματα, είναι ανάλογες με τους διαδοχικούς μονούς αριθμούς.

Το σχήμα (b) δείχνει ένα σώμα να κυλά επί ενός κεκλιμένου επιπέδου κλίσης 45° . Οι διαδοχικές αποστάσεις έχουν μεταξύ τους την ίδια σχέση με πριν, με τη διαφορά ότι είναι μικρότερες κατά τον παράγοντα 0,7, το συνημίτονο της γωνίας 45° .

Στην τελευταία περίπτωση (σχήμα c), η πτώση περιορίζεται περισσότερο με την κλίση του επιπέδου να αυξάνεται στις 60° . Οι αποστάσεις εδώ είναι μικρότερες κατά 0,5.

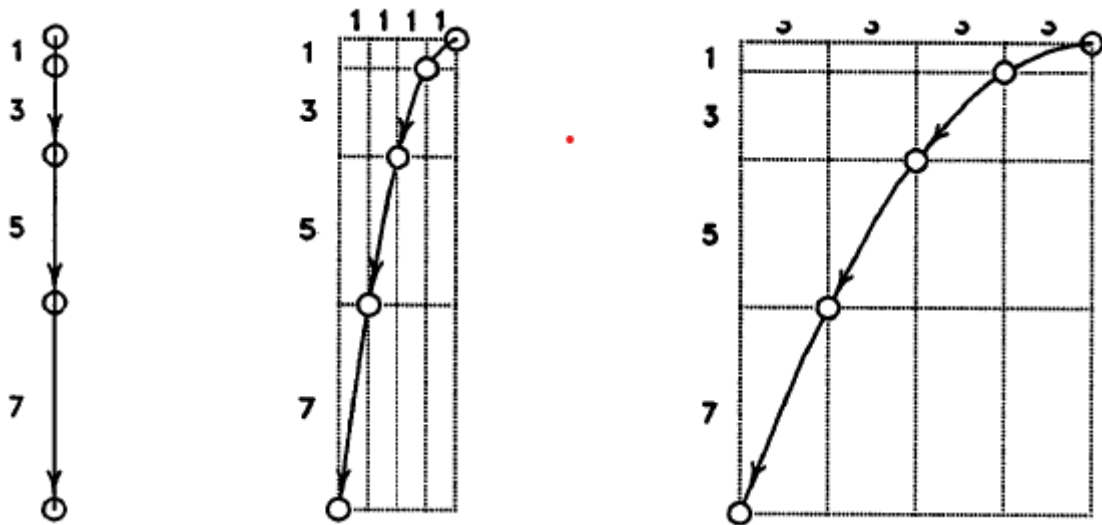
Μπορεί, πλέον, το αρχικό σχήμα να ειπωθεί με παρόμοιο τρόπο με τα άλλα δύο. Το κεκλιμένο επίπεδο στην πρώτη περίπτωση έχει γωνία 0° .



Εικόνα 8: (a) Κάθετη περίπτωση. (b) Κλίση 45 μοιρών. (c) Κλίση 60 μοιρών

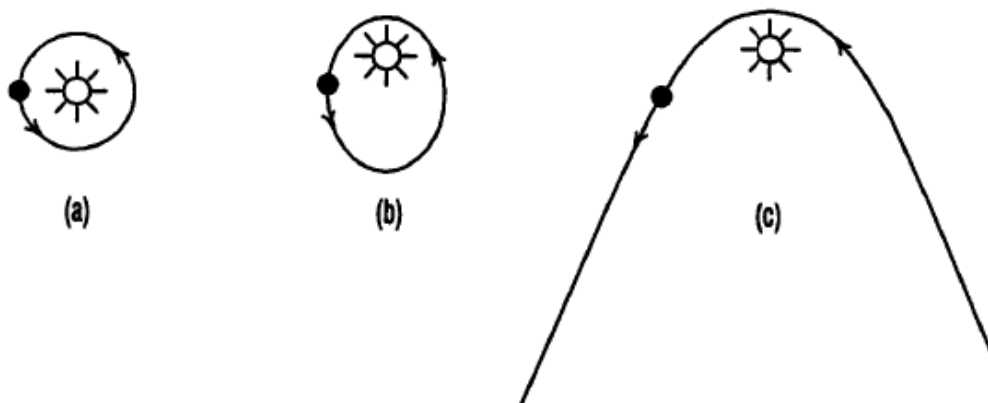
Μια παραλλαγή του βασικού θέματος παρουσιάζεται στο επόμενο σχήμα. Το σχήμα (b) διαφέρει από το (a) στο ότι το σώμα έχει τώρα μια ομοιόμορφη οριζόντια κίνηση. Ο συνδυασμός της κατακόρυφης και οριζόντιας κίνησης μας δίνει αυτή τη χαρακτηριστική

παραβολική τροχιά. Το πλάτος της τροχιάς εξαρτάται από το μέγεθος της αρχικής ομοιόμορφης οριζόντιας κίνησης όπως φαίνεται στο σχήμα (c). Το πρώτο σχήμα φαίνεται πλέον όμοιο με τα υπόλοιπα, εκτός του ότι η οριζόντια ταχύτητά του είναι μηδενική.



Εικόνα 9: Ελεύθερη πτώση: (a) κάθετη περίπτωση. (b) Μικρή συνιστώσα ομοιόμορφης οριζόντιας κίνησης. (c) Μεγαλύτερη συνιστώσα ομοιόμορφη οριζόντιας κίνησης

Με τα επόμενα διαγράμματα ο μαθητής μπορεί να εισαχθεί στην ανοδική κίνηση υπό σταθερή επιβράδυνση. Τώρα το φάσμα των δυνατοτήτων διευρύνεται, συμπεριλαμβάνοντας την ελεύθερη πτώση υπό την επίδραση μιας ανομοιόμορφης βαρυτικής δύναμης, όπως συμβαίνει με τους πλανήτες σε τροχιά γύρω από τον ήλιο.



Εικόνα 10: Σώματα που πέφτουν ελεύθερα σε ανομοιόμορφο βαρυτικό πεδίο. (a) κυκλική περίπτωση: η επιτάχυνση προς τον ήλιο είναι σταθερή. (b) Ελλειπτική περίπτωση: η επιτάχυνση δεν είναι σταθερή αλλά κυκλική. (c) Υπερβολική περίπτωση: η επιτάχυνση προς τον ήλιο δεν είναι ούτε σταθερή ούτε κυκλική.

Ο Churchland διαπιστώνει ότι, όσον αφορά τη μάθηση μιας θεωρίας, «είναι λιγότερο ζήτημα απομνημόνευσης ενός συνόλου προτάσεων και περισσότερο ζήτημα εξοικείωσης με μια οικογένεια παραδειγμάτων ή πρωτοτυπικών αιτιωδών διαδικασιών που προσδιορίζονται σε αυτήν τη θεωρία» (Churchland, 1994, p.274). Η παράλειψη των σχετικών εξισώσεων που περιγράφουν τα παραπάνω φαινόμενα, από τη μεριά του

Churchland, δεν είναι τυχαία. Ο Kuhn επιχειρεί να δείξει ότι η κατανόηση αυτών των παραδειγμάτων και παρόμοιων που σχετίζονται με αυτά, είναι σημαντικότερη στη συνολική κατανόηση της θεωρίας, εκ μέρους του μαθητή, από ότι μια στείρα παράθεση εξισώσεων.

Συμπεραίνουμε, λοιπόν, ότι είναι η κατανόηση ορισμένων παραδειγματικών τύπων καταστάσεων και διαδικασιών, καθώς και των δυνατών παραλλαγών των βασικών τους θεμάτων, το σημαντικότερο εργαλείο για την κατανόηση μιας θεωρίας. Μορφές ρητών προτασιακών αναπαραστάσεων όπως οι εξισώσεις, είναι συχνά σημαντικές, αλλά «η ικανότητα χρήσης τους είναι απλώς μια πλευρά ενός ευρύτερου πλαισίου δεξιοτήτων (αντιληπτικών, εννοιολογικών, ερμηνευτικών, αναλογικών, μετασχηματιστικών και χειρισμού)· ενός πλαισίου που αποτελεί τον αληθινό μοχλό κατανόησης» (Churchland, 1994, p.275).

4.5.2 Ο ρόλος του «παραδείγματος» στη νευροϋπολογιστική θεωρία του Churchland

Ο Churchland οικειοποιείται τις θέσεις και τα συμπεράσματα στα οποία καταλήγει ο Kuhn, για να τα μεταφέρει στο πεδίο των νευρωνικών δικτύων. Έτσι, η έννοια του πρωτότυπου διανύσματος αντιστοιχίζεται με το «παραδείγμα», το οποίο αποτελεί για τον Kuhn μια αντικειμενική κατασκευή. Το φάσμα των ικανοτήτων που αναπτύσσονται κατά την εξάσκηση ενός νευρωνικού δικτύου στην αντίστοιχη ιεραρχία εσωτερικών πρωτοτυπικών διανυσμάτων, σχετίζεται με το φάσμα των ικανοτήτων που εμπλέκονται στην επίλυση προβλημάτων μέσω της κατανόησης μιας οικογένειας παραδειγμάτων. Η κατανόηση της θεωρίας, επομένως, συνίσταται «σε μια οικογένεια ικανοτήτων που κωδικοποιούνται ως μια ιεραρχία πρωτοτύπων και πρωτοτυπικών ακολουθιών στο χώρο των νευρωνικών ενεργοποιήσεων του εγκεφάλου», κι όχι στην αποδοχή και το χειρισμό ενός δεδομένου συνόλου προτάσεων (Churchland, 1994, p.277).

Για τον Churchland, η επιστημονική γνωσιακή δραστηριότητα είναι όμοια σε είδος με την καθημερινή, συνηθισμένη γνωσιακή δραστηριότητα. Το επιστημονικό γινώσκειν ξεχωρίζει από τη σχετική καινοτομία του, τη φιλοδοξία του, τις θεσμοθετημένες διαδικασίες και την πραγματολογική ισχύ του. Το συμπέρασμα αυτό παρέχει την προοπτική για μια μεγάλη γνωσιακή ανάπτυξη ολόκληρων κοινωνιών. Ο Churchland φιλοδοξεί οι καθημερινές έννοιες που χρησιμοποιούμε στις πρακτικές μας δραστηριότητες να αντικατασταθούν από τις αντίστοιχες ισχυρότερες επιστημονικές έννοιες, έτσι ώστε οι άνθρωποι να αποκτήσουν μια επαρκέστερη γνωσιακή κατανόηση και έλεγχο του κόσμου. «Το μυστηριώδες θεωρητικό πλαίσιο του σήμερα μπορεί να γίνει η καθημερινή συνηθισμένη γνώση του αύριο», όπως σημειώνει (Churchland, 1994, p.278). Σκοπός του τελικά, είναι να γίνει η επιστημονική γνώση κτήμα ολόκληρης της κοινωνίας.

Περίοπτη θέση επιφυλάσσει και για τη δημιουργικότητα, έννοια την οποία μπορούμε να εξετάσουμε διεισδυτικότερα από την προοπτική των νευρωνικών δικτύων. Όπως και η συνείδηση, η δημιουργικότητα αποτελεί ένα, μάλλον, πολυδιάστατο φαινόμενο. Στις μεγάλες επιστημονικές ανακαλύψεις, όμως, κυριαρχεί μια χαρακτηριστική της διάσταση. «Ο Αριστοτέλης είδε τον ουρανό ως μια περιστρεφόμενη σφαίρα, ο Καρτέσιος είδε το

ηλιακό σύστημα ως μια δίνη από διαφανή ύλη, ο Νεύτωνας είδε το φεγγάρι και τους πλανήτες ως σώματα σε ελεύθερη πτώση με μια εφαπτόμενη αδρανειακή κίνηση». Παρατηρούμε, δηλαδή, ότι οι επιστήμονες χρησιμοποιούν εργαλεία από το εννοιολογικό ρεπερτόριο που ήδη διαθέτουν ώστε να ερμηνεύσουν ένα προβληματικό φαινόμενο ως μια μη αναμενόμενη ή ασυνήθιστη περίπτωση ενός πρωτοτυπικού προτύπου (Churchland, 1994, p.278).

Ο Αριστοτέλης, ο Καρτέσιος, ο Νεύτωνας κατάφεραν να εξερευνήσουν διαφορετικές ενεργοποιητικές δυνατότητες μέσω των επανατροφοδοτικών τους δικτύων, δυνατότητες ήδη παρούσες στην ιεραρχία των κατανομών τους. Αυτές οι πρωτοτυπικές δυνατότητες συνυπάρχουν σε ένα χώρο ομοιοτήτων με άλλες μη τυπικές. Η επανατροφοδοτική δραστηριότητα που φτάνει σε αυτόν τον πληθυσμό των νευρώνων μπορεί να μεταβάλει τις γνωσιακές αποκρίσεις του, σε μια προβληματική εισροή, με τον ένα ή τον άλλο τρόπο, με την ελπίδα να ενεργοποιηθεί κάποιο διάνυσμα που βρίσκεται κοντά σε ένα οικείο πρωτότυπο, ώστε η λογική συνοχή του προβληματικού φαινομένου να αποκατασταθεί.

4.5.3 Η εννοιολογική επέκταση στο έργο του Αριστείδη Μπαλτά

Στο βιβλίο του «Ξεφλουδίζοντας πατάτες ή λειαινώντας φακούς», ο Αριστείδης Μπαλτάς αναζητά συνδέσεις και δρόμους επικοινωνίας ανάμεσα σε δύο από τους επιδραστικότερους φιλοσόφους των τελευταίων αιώνων, τον Ludwig Wittgenstein και τον Baruch Spinoza. Ο Μπαλτάς διερευνά την *προοπτική της ριζικής εμμέλειας* με τη μορφή αντιστικτικής ανάγνωσης της «Ηθικής» και του «Tractatus Logico-Philosophicus». Η βασική υπόθεση εργασίας την οποία υιοθετεί ο Μπαλτάς εδράζεται στο ότι ο μεν Spinoza αποσκοπεί να εξαλείψει τη θέση έξωθεν του Θεού-Δημιουργού και συγκεκριμένα του ανθρωπομορφικού, με τη πλήρη αυστηρότητα της γεωμετρικής τάξης, ταυτίζοντας το Θεό με τη Φύση, ενώ ο Wittgenstein την έξωθεν δεσπόμενη θέση της λογικής επί της σκέψης και της γλώσσας και τελικά επί του κόσμου, επίσης με απόλυτη – λογική αυτή τη φορά – αυστηρότητα. Η λογική είναι μέσα στη γλώσσα για τον Wittgenstein ακριβώς όπως ο Θεός είναι μέσα στον κόσμο για τον Spinoza. Ο Spinoza αποσκοπεί να αποδομήσει την ιδέα ότι μπορούμε να δικαιώσουμε τον Θεό μόνον αν τον αναβιβάσουμε υπεράνω και έξωθεν του κόσμου. Ο Wittgenstein αποσκοπεί να αποδομήσει την ιδέα ότι μπορούμε να δικαιώσουμε τη λογική μόνον αν την αναβιβάσουμε σε κάτι πρότερο των συνθηκών που καθιστούν τις εκάστοτε εμπειρικές προτάσεις αληθείς ή ψευδείς. Το βιβλίο επιχειρηματολογεί ότι αυτή η αποδόμηση παραδοσιακών παραδοχών οδηγεί και στις δύο περιπτώσεις σε μια θέση σύμφωνα με την οποία η φιλοσοφία, όπως νοείται παραδοσιακά, είναι αδύνατη και ταυτόχρονα στην απόρριψη κάθε σκοπιάς έξω από τον κόσμο, τη σκέψη και τη γλώσσα.

Σε διάφορα σημεία του βιβλίου γίνεται λόγος για την «διεύρυνση του διαθέσιμου γραμματικού χώρου», μια έννοια εξέχουσας σημασίας στο έργο του Μπαλτά. Σε μια προσπάθεια διάκρισης εννοιολογικών και λογικών δυνατοτήτων, ο συγγραφέας επικαλείται ένα παράδειγμα από την ιστορία της επιστήμης. Στη νευτώνεια φυσική, η έννοια «κύμα» οριζόταν ως διάδοση τοπικών διαταραχών ενός μέσου, ήταν, με άλλα λόγια, «αναλυτικά συνδεδεμένη» με την έννοια «μέσο». Άρα η πρόταση «ένα κύμα διαδίδεται στο κενό»,

συνιστά λογική αντίφαση, καθώς το κενό αναφέρεται στην απουσία οποιουδήποτε μέσου. Η Ειδική Θεωρία της Σχετικότητας, παρ'όλ'αυτά, θα αλλάξει ριζικά τη θεώρηση του ζητήματος με τη θεωρητική και εμπειρική τεκμηρίωση της ύπαρξης ηλεκτρομαγνητικών κυμάτων, τα οποία διαδίδονται στο κενό, χωρίς την ανάγκη κάποιου μέσου διάδοσης. «Διαταραχές του τίποτα που διαδίδονται στην έκταση του τίποτα», σημειώνει ο Μπαλτάς, «μια νέα εννοιολογική δυνατότητα, πειραματικά επικυρωμένη και θεωρητικά συνεπής» θα επιβληθεί με εμφατικό τρόπο ξεπερνώντας την προϋπάρχουσα λογική αντίφαση. Υπό αυτή την έννοια, μπορεί να γίνει λόγος για διεύρυνση και μετασχηματισμό του διαθέσιμου γραμματικού χώρου. Αυτή η διεύρυνση μπορεί να συμπεριλάβει εννοιολογικά τόσο το νέο ορισμό (του κύματος εν προκειμένω) όσο και τον ξεπερασμένο, ειδωμένο πλέον υπό τους όρους του διευρυμένου γραμματικού χώρου. Αναφερόμενος σε αυτήν την τομή σαν *συμβάν*, ο Μπαλτάς υποστηρίζει ότι «μετά το συμβάν, η λογική αποκαθίσταται», καθώς η προηγούμενη λογική αντίφαση ερμηνεύεται τώρα σαν «στενότητα γραμματικού χώρου, ως αδυναμία του παλιού γραμματικού πλαισίου να στεγάσει τη νέα έννοια» (Μπαλτάς, 2022).

Σημαντική λειτουργία των εννοιολογικών επαναστάσεων ή των αλλαγών παραδείγματος αποτελεί αυτή η διεύρυνση του διαθέσιμου γραμματικού χώρου με τρόπους που, ενώ είναι λογικά αδύνατο να συλληφθούν πριν την επανάσταση, εκ των υστέρων επιβεβαιώνεται ότι η λογική παραμένει άθικτη. Οι λογικές δυνατότητες, εν αντιθέσει με τις γραμματικές, παραμένουν αμετάβλητες κατά τη διαδοχή των επιστημονικών μοντέλων. Η ριζική αλλαγή επιστημονικού παραδείγματος, διακόπτοντας την ομαλή ροή της ιστορίας της επιστήμης, αποτελεί μείζον συμβάν που δυναμιτίζει τα θεμέλια της φιλοσοφίας. Μια τέτοια ριζική αλλαγή εισάγει στον κόσμο ένα νέο κεκτημένο το οποίο ήταν, μέχρι πρότινος, απρόσιτο στη φιλοσοφία.

4.6 Το πείραμα του Κινέζικου Δωματίου και τα qualia

Οι απόψεις του Paul Churchland - και άλλων προγενέστερων συγγενικών του προσεγγίσεων σχετικά με τη φύση της συνείδησης και του νου - έχουν προκαλέσει σειρά αντιπαραθέσεων εντός της φιλοσοφικής και επιστημονικής κοινότητας. Στις επόμενες παραγράφους θα διερευνηθούν οι πιο σημαντικές από αυτές.

Ο σκεπτικισμός γύρω από την προοπτική της μηχανικής νοημοσύνης εστιάζει σε τρία κυρίως ζητήματα. Το πρώτο αναφέρεται στο ζήτημα του νοήματος ή σημασιολογικού περιεχομένου των ανθρώπινων νοητικών καταστάσεων όπως οι σκέψεις, οι επιθυμίες και οι πεποιθήσεις. Το δεύτερο αφορά τη δυνατότητα ενός υπολογιστή να αναπαραγάγει πλήρως την ανθρώπινη ικανότητα στα μαθηματικά, ενώ το τελευταίο κατά σειρά ζήτημα ασχολείται με την προβληματική του ποιοτικού χαρακτήρα της συνειδητής εμπειρίας.

4.6.1 Ο Searle μέσα στο δωμάτιο

Ο John Searle επινόησε ένα πείραμα σκέψης για να δείξει ότι η νοημοσύνη δεν μπορεί να ταυτίζεται με τη λειτουργία ενός προγράμματος υπολογιστή, ανεξάρτητα από το πόσο περίπλοκο ή έξυπνο μπορεί να φαίνεται.

Φανταστείτε ότι ένα αγγλόφωνο άτομο, το οποίο δεν γνωρίζει καθόλου κινέζικα, είναι κλειδωμένο σε ένα δωμάτιο. Ένα άτομο έξω από το δωμάτιο κάνει ερωτήσεις στα κινέζικα, μέσω γραπτών μηνυμάτων, και το άτομο στο δωμάτιο απαντά επίσης χρησιμοποιώντας κινέζικα σύμβολα. Στο δωμάτιο υπάρχει ένα βιβλίο γραμμένο στα αγγλικά, με ένα σύνολο οδηγιών για το πώς να χειρίζεται ακολουθίες κινεζικών συμβόλων. Η εφαρμογή των οδηγιών εξαρτάται εξ ολοκλήρου από τη μορφή των συμβόλων και είναι ανεξάρτητη από τη σημασία τους. Υπάρχει επίσης ένα βιβλίο, μια βάση δεδομένων όπου ο άνθρωπος μπορεί να βρει τα κινεζικά σύμβολα που χρησιμοποιεί στα εξερχόμενα μηνύματα. Έτσι, ο άνθρωπος έξω λαμβάνει απαντήσεις στα κινέζικα και πιστεύει ότι ο άνθρωπος μέσα ξέρει να μιλάει κινέζικα. Όμως, είναι σαφές ότι ο άνθρωπος μέσα δεν καταλαβαίνει καθόλου κινέζικα. Το πρόγραμμα επιτρέπει στον άνθρωπο μέσα στο δωμάτιο να περάσει το τεστ Turing για την κατανόηση των κινεζικών, χωρίς να καταλαβαίνει ούτε μια λέξη κινεζικών (Searle, *Minds, brains, and programs*, 1980).

Αυτό που κάνει ο άνθρωπος μέσα στο ‘κινεζικό δωμάτιο’ είναι αυτό που θα έκανε ένας υπολογιστής, δηλαδή χειρισμός συντακτικών συμβόλων. Το πρόγραμμα στο δωμάτιο είναι το βιβλίο των εντολών, η μνήμη είναι το σύνολο των δεδομένων και ο άνθρωπος παίζει το ρόλο της CPU. Επομένως, όπως ακριβώς ο άνθρωπος στο δωμάτιο που δεν κατανοεί καθόλου τα κινεζικά, έτσι και κανένας υπολογιστής, απλώς και μόνο ακολουθώντας ένα πρόγραμμα, δεν καταλήγει να κατανοεί πραγματικά τα κινεζικά. Καμία πραγματική κατανόηση δεν μπορεί να παραχθεί μόνο από το συντακτικό. Η πραγματική κατανόηση των συμβόλων προϋποθέτει τη σημασιολογία, τη γνώση του τι σημαίνουν αυτά τα σύμβολα, τι αντιπροσωπεύουν. Έτσι, το τεστ Turing είναι ανεπαρκές.

Το ‘κινεζικό δωμάτιο’ του Searle στρέφεται κυρίως κατά του λειτουργισμού. Ο ανθρώπινος νους δεν είναι απλώς ένα υπολογιστικό σύστημα που μοιάζει με υπολογιστή και έτσι, η κεντρική ιδέα αυτού που προηγουμένως ονομάσαμε ισχυρή τεχνητή νοημοσύνη, είναι λανθασμένη. Δεν μπορούμε να πάμε από το συντακτικό στη σημασιολογία. Όταν δίνεται μια είσοδος από 1 και 0 σε έναν υπολογιστή, αυτός θα παράγει μια άλλη ακολουθία από 1 και 0, ανεξάρτητα από το αν εμείς, με κάποιο τρόπο, δώσουμε ένα νόημα σε αυτή την ακολουθία, είτε σημαίνει «ένα φλιτζάνι τσάι» είτε «αριθμός χωρών στον κόσμο». Ο υπολογιστής θα παρήγαγε την ακολουθία, ακόμη και αν δεν σήμαινε τίποτα. Με την αναγωγή των νοητικών καταστάσεων στον συντακτικό χειρισμό συμβόλων, αποτυγχάνουμε να εξηγήσουμε την πιο σημαντική ιδιότητα της νοημοσύνης: το νόημα και την προθετικότητα των νοητικών καταστάσεων. Ο Searle το εξήγησε με έναν ενδεικτικό τρόπο:

«Τα υπολογιστικά μοντέλα της συνείδησης δεν επαρκούν από μόνα τους για τη συνείδηση. Το υπολογιστικό μοντέλο για τη συνείδηση στέκεται στη συνείδηση με τον ίδιο τρόπο που το υπολογιστικό μοντέλο οποιουδήποτε πράγματος στέκεται στον τομέα που μοντελοποιείται. Κανείς δεν υποθέτει ότι το υπολογιστικό μοντέλο των βροχοπτώσεων στο

Λονδίνο θα μας αφήσει όλους βρεγμένους. Κάνουν όμως το λάθος να υποθέσουν ότι το υπολογιστικό μοντέλο της συνείδησης είναι κατά κάποιον τρόπο συνειδητό. Πρόκειται για το ίδιο λάθος και στις δύο περιπτώσεις» (Searle, *Consciousness and Language*, 2002).

Ο John Searle υιοθετεί μια πολυσυζητημένη άποψη σχετικά με τη φύση του νοήματος ή του σημασιολογικού περιεχομένου. Ισχυρίζεται ότι το γνήσιο νόημα είναι εγγενές στις ανθρώπινες καταστάσεις της συνείδησης, αλλά απουσιάζει από αυτές ενός ηλεκτρονικού υπολογιστή. Οι καταστάσεις του υπολογιστή έχουν στην καλύτερη των περιπτώσεων, ένα δευτερεύον ή οιονεί είδος νοήματος, ένα είδος που αναδύεται κατ'αρχήν μόνο και μόνο επειδή οι άνθρωποι το βρίσκουμε χρήσιμο και βολικό να ερμηνεύουμε τις καταστάσεις του ως αναπαραστάσεις κάποιων αριθμών ή προτάσεων. Ο Searle υποστηρίζει ότι οι καταστάσεις του δεν έχουν εγγενές νόημα, ακριβώς όπως δεν έχουν και οι διάφοροι σχηματισμοί από τις χάντρες ενός άβακα. Κανενός είδους χειρισμός των χαντρών δε θα αποδώσει στους σχηματισμούς τους κάποιο εγγενές νόημα. Όμοια, κανενός είδους χειρισμός των καταστάσεων ενός υπολογιστή δεν θα τους αποδώσει εγγενές νόημα. Καταλήγει έτσι στο συμπέρασμα ότι ο προγραμματισμένος χειρισμός φυσικών πραγμάτων δεν έχει τη δύναμη να παραγάγει εγγενές νόημα.

Για τον Churchland, ο Searle υιοθετεί μια υβριδική θέση ανάμεσα στις παλιές δυϊστικές και τις πιο σύγχρονες νευροεπιστημονικές απόψεις. Η στάση του διαφέρει από προηγούμενες δυϊστικές και αντι-αναγωγιστικές προσεγγίσεις, καθώς δέχεται ότι οι αισθήσεις, οι σκέψεις και γενικότερα τα νοητικά φαινόμενα αποτελούν καταστάσεις ή χαρακτηριστικά του εγκεφάλου. Ωστόσο, υποστηρίζει ότι οι νοητικές καταστάσεις δεν είναι καθαυτές φυσικές καταστάσεις του εγκεφάλου. Αποτελούν μια διαφορετική κατηγορία φαινομένων με ιδιάζουσες ιδιότητες και τρόπους συμπεριφοράς (όπως νόημα και *προθετικότητα*), ενώ δε δύναται να ταυτιστούν ή να αναχθούν σε απλά φυσικά φαινόμενα. Η σχέση νοητικών και φυσικών καταστάσεων του εγκεφάλου έχει αιτιώδη χαρακτήρα για τον Searle· «οι καταστάσεις του εγκεφάλου είναι η αιτία των νοητικών καταστάσεων, και αντίστροφα» (Churchland, 1994, p.203).

Ο Churchland χαρακτηρίζει ως «συντηρητική σύγχρονη», τη θέση του Searle σχετικά με τη φύση των νοητικών φαινομένων. 'Συντηρητική' λόγω της αμετακίνητης στάσης της απέναντι στην ανεξάρτητη πραγματικότητα και στο διακριτό μεταφυσικό χαρακτήρα των νοητικών καταστάσεων, και 'σύγχρονη' γιατί μεταφέρει τα νοητικά φαινόμενα και «τα εγκαθιστά ως (μη φυσικά) χαρακτηριστικά στον εγκεφαλο», κατάλληλα για επιστημονική μελέτη (Churchland, 1994, p.203). Ο Searle, σύμφωνα τον Churchland, αναγκάζεται να υιοθετήσει αυτήν την επαμφοτερίζουσα θέση από τη στιγμή που δέχεται τα επιχειρήματα για τις υποκειμενικές ιδιότητες (βλ. Nagel κλπ). Όμως για τον Churchland, αυτά τα επιχειρήματα δεν αποδεικνύουν ότι οι νοητικές καταστάσεις πρέπει να μην ταυτίζονται με τις φυσικές καταστάσεις του εγκεφάλου. «Δείχνουν μόνο ότι ο καθένας μας έχει το δικό του, προεπιστημονικό τρόπο γνώσης της εμφάνισης και του χαρακτήρα των ίδιων των εσωτερικών καταστάσεών του» (Churchland, 1994, p.204).

Σύμφωνα με τον Churchland, η σημασιολογική θεωρία του Searle είναι αμφίβολη. Είναι αμφίβολο το αν υπάρχει κάποιο «εγγενές νόημα», ένα νόημα το οποίο έχει μια

κατάσταση ανεξάρτητα από τις σχέσεις που έχει ή δεν έχει με άλλες καταστάσεις και με τον εξωτερικό κόσμο. Εάν το νόημα μπορούσε να αποσυνδεθεί με αυτόν τον τρόπο και να καταστεί εγγενές, τότε θα έπρεπε να είναι δυνατό ένα φυσικό σύστημα να έχει μία και μόνο μία εσωτερική κατάσταση. Ο Churchland αναφέρει χαρακτηριστικά, «είναι σαν να επιμένεις ότι κάποιο πρόσωπο θα μπορούσε να περιγραφεί ορθά ως 'ο νεότερος γερουσιαστής από την Καλιφόρνια', χωρίς να υπάρχει ένα πολύ μεγάλο δίκτυο άλλων αιρετών θέσεων, νομοθετικών θεσμών, υπουργείων, πολιτικών κομμάτων, εκλογικών διαδικασιών, εγγεγραμμένων ψηφοφόρων και κάποια όμορφη περιοχή που τη λένε Καλιφόρνια» (Churchland, 1994, p.245).

4.6.2 Οι αισθητηριακές ποιότητες εντός του λειτουργισμού

Το πρόβλημα με τις αισθητηριακές ποιότητες (qualia), σύμφωνα με τον Churchland, έγκειται στην εύρεση ενός εύλογου πεδίου για αυτές μέσα σε ένα αμιγώς λειτουργιστικό πλαίσιο. Ο Churchland αναφέρει την περίπτωση του Ned Block, ενός φιλοσόφου από το MIT, ανήσυχου για ένα φανερό πρόβλημα με το λειτουργισμό, την κυρίαρχη μορφή του υλισμού στον καιρό του. Οι λειτουργιστές υποστήριζαν ότι η ουσία της ενσυνείδητης νοημοσύνης βρίσκεται στο λογισμικό (software), στο αφηρημένο δηλαδή υπολογιστικό πρόγραμμα, στο σύνολο των αλγοριθμικών διεργασιών, οι οποίες εκτελούνται σε κάθε φυσιολογικό άνθρωπο από το βιολογικό υλισμικό (hardware) του. Αυτή ήταν μια σχεδόν καθολικά αποδεκτή θέση από τους ερευνητές της Τεχνητής Νοημοσύνης τον καιρό του Block. Αντιμετώπιζαν το πρόβλημα της δημιουργίας αληθινής νοημοσύνης σε μια μηχανή απλά ως ένα «ζήτημα γραφής του προγράμματος που 'εκτελείται' στους φυσιολογικούς ανθρώπους ή ενός προγράμματος ισοδύναμου ως προς την είσοδο-έξοδο, με αυτό των ανθρώπων» (Churchland, 1994, p.250). Το είδος της μηχανής στην οποία θα έτρεχε το πρόγραμμα ήταν, το πολύ, δευτερεύουσας σημασίας. Κι αυτή είναι η κλασική θέση την οποία προσπαθεί να καταρρίψει στο συγκεκριμένο έργο του ο Churchland.

Ο συγγραφέας, ακολουθώντας τη συλλογιστική του Block, μας καλεί να φανταστούμε την ύπαρξη αυτού του ασυνήθιστου ανθρώπινου προγράμματος. Για την εκτέλεση αυτού του προγράμματος οργανώνεται ένα αρκετά μεγάλο πλήθος ατόμων, εφοδιασμένα το καθένα με ειδικές κάρτες, ακολουθώντας κανόνες και αλληλεπιδρώντας με άλλα άτομα, με τα οποία θα έχει τον έλεγχο, μέσω ραδιοκυμάτων, ενός και μοναδικού ρομπότ. Αν και η εκτέλεση θα ήταν υπερβολικά αργή, θα μπορούσε να εκτελείται η ίδια αχανής συνάρτηση εισόδου-εξόδου που εκτελεί ένας άνθρωπος. Ακόμα και σε αυτήν την περίπτωση, είναι αρκετά αμφίβολο ότι ένα τέτοιο ευρύ κοινωνικό σύστημα, θεωρούμενο ως ένα πράγμα, θα μπορούσε να έχει αισθήσεις με εσωτερικά ποιοτικά χαρακτηριστικά του ίδιου είδους με αυτά που διαθέτει ένας άνθρωπος. Αν και όλες οι αλγοριθμικές συνθήκες φαίνεται να ικανοποιούνται, οι αισθητηριακές ποιότητες απουσιάζουν από το σύστημα. Αυτή η στενά λειτουργιστική θέση «παραλείπει κάτι σημαντικό σχετικά με τη φύση της ανθρώπινης νοημοσύνης» (Churchland, 1994, p.251).

Το σφάλμα του λειτουργισμού ήταν ότι αδιαφόρησε για το ποιες ακριβώς διεργασίες λαμβάνουν χώρα μέσα στο άτομο· αρκούσαν απλά στην εύρεση της σωστής συνάρτησης εισόδου-εξόδου, με αποτέλεσμα η εσωτερική εικόνα του γνωσιακού πλάσματος - που αποτελεί ο άνθρωπος - να αποδεικνύεται μερική και αποσπασματική. Μάλιστα, όταν τις έλαβε υπόψη, «η υπόθεσή του σχετικά με τη φύση αυτών των διεργασιών ήταν λανθασμένη: τις εξέλαβε ως αλγοριθμικές στον πυρήνα τους» (Churchland, 1994, p.251).

Σήμερα, η προοπτική του κλασικού λειτουργισμού θεωρείται παρωχημένη, εφόσον πλέον γνωρίζουμε ότι οι φυσικές διεργασίες που συμβαίνουν εντός μας δεν εκτελούν απλά ένα πρόγραμμα, αλλά έχουν και σημασία. Οι περισσότερες θεμελιώδεις ιδιότητες της ανθρώπινης και ζωικής αντίληψης εμφανίζονται εξαιτίας της ιδιαίτερης φυσικής οργάνωσης του νευρικού συστήματος, του ιδιαίτερου τρόπου κωδικοποίησης αυτών των πληροφοριών και των φυσικά κατανεμημένων μέσων τα οποία μετασχηματίζουν τις πληροφορίες (Churchland, 1994, p.251).

Ο Churchland θεωρεί το επιχείρημα για τα qualia, ως επιχείρημα ενάντια σε μια πολύ ειδική εκδοχή του υλισμού, την εκδοχή σύμφωνα με την οποία η γνωσιακή δραστηριότητα είναι ένας χειρισμός διακριτών φυσικών συμβόλων καθοδηγούμενος από κανόνες. Η νευροϋπολογιστική θεώρηση μπορεί και εδώ να προσφέρει λύσεις, καθώς ο χαρακτήρας της αισθητηριακής διανυσματικής κωδικοποίησης και οι δομές μέσα σε ένα δεδομένο αισθητηριακό διανυσματικό χώρο παρέχουν σημαντικά εργαλεία για μια αρχική κατανόηση του προβλήματος.

4.6.3 Πώς είναι να είσαι;

Το 1974 ο Αμερικανός φιλόσοφος Thomas Nagel δημοσίευσε το δημοφιλές του άρθρο με τίτλο «Πώς είναι να είσαι νυχτερίδα;» (What is it like to be a bat) επιχειρώντας να αναδείξει το πρόβλημα του βιώματος της υποκειμενικής εμπειρίας και της αντίληψης της πραγματικότητας, σε σχέση με το πρόβλημα της αναγωγής της «συνείδησης» σε αμιγώς υλική βάση. Ο Nagel υποστήριξε ότι ένας νοήμων οργανισμός βρίσκεται σε αυτή τη νοητική κατάσταση που ονομάζουμε συνείδηση, όταν έχει αυτήν την ιδιαίτερη αίσθηση του «πώς είναι να είσαι» αυτός ο συγκεκριμένος οργανισμός. Το νοητικό πείραμα που παρουσιάζεται στο άρθρο, λαμβάνει χώρα εντός του εγκεφάλου μιας νυχτερίδας, ενός θηλαστικού με αρκετά διαφορετικό εύρος δραστηριοτήτων και αισθητηριακό μηχανισμό από τον ανθρώπινο.

Ο συγγραφέας ισχυρίζεται ότι η γνώση μας για τη νευροανατομία του εγκεφάλου της νυχτερίδας και για τη νευροφυσιολογία της αισθητηριακής της δραστηριότητας, δε λέει τίποτα για το «πώς είναι το να έχεις» τις αισθητηριακές εμπειρίες της νυχτερίδας. Οι άνθρωποι αγνοούμε το πώς είναι αυτές οι εμπειρίες από την προοπτική του πρώτου προσώπου της νυχτερίδας, από την προοπτική που έχει η νυχτερίδα ως το μοναδικό τους υποκείμενο. Το κενό μεταξύ φυσικής πραγματικότητας του βιολογικού εγκεφάλου και της ψυχολογικής πραγματικότητας της πρωτοπρόσωπης συνειδητής εμπειρίας καθίσταται εμφανές. Ο Nagel αντικρούοντας τις αναγωγιστικές αντιλήψεις, συμπεραίνει ότι τα φαινόμενα της συνείδησης δεν επιδέχονται μια καθαρά φυσική εξήγηση (Nagel, 1974).

Ο Churchland εξετάζει λεπτομερώς το επιχείρημα του Nagel, από τη σκοπιά των νευρωνικών δικτύων. Η νυχτερίδα, όπως και κάθε άλλος φυσικός οργανισμός, έχει μια ιδιάζουσα πρόσβαση στις αισθήσεις της, πρόσβαση που δε διαθέτει ένας εξωτερικός παρατηρητής. Αυτό συμβαίνει γιατί το κάθε πλάσμα απολαμβάνει ένα μοναδικό σύνολο αιτιωδών συνδέσεων με την αισθητηριακή δραστηριότητα του εγκεφάλου του και του νευρικού του συστήματος. Κάθε άτομο λαμβάνει πληροφορίες σχετικά με τις συνεχείς μεταβολές της ίδιας της αισθητηριακής δραστηριότητάς του μέσω ενός ειδικού συνόλου νευρωνικών οδών, που μόνο αυτό διαθέτει. Τα υπόλοιπα άτομα έχουν ποιοτικά όμοιες νευρικές οδούς, αλλά σχηματίζουν συνδέσεις μεταξύ των δικών τους αισθητηριακών δραστηριοτήτων και των δικών τους εγκεφάλων.

«Κάθε πλάσμα έχει έναν τρόπο γνώσης των δικών του αισθητηριακών καταστάσεων που δεν έχει κανένα άλλο πλάσμα», θα συμφωνήσει ο Churchland. Εναλλακτικοί τρόποι γνώσης των προσωπικών αισθητηριακών καταστάσεων κάποιου ατόμου, μπορεί να είναι η συμπεριφορά του, η εξέταση PET και η τοποθέτηση ηλεκτροδίων στον εγκέφαλο. Η γνώση αυτών των καταστάσεων μέσω των ατομικών πληροφοριακών οδών παραμένει, ωστόσο, στην κατοχή του ατόμου που διαθέτει τις νευρικές οδούς. Αυτές «είναι οι οδοί που συνθέτουν τον εγκέφαλο και το νευρικό σου σύστημα, οι οδοί που συνιστούν τη δική σου ιεραρχία νευρωνικών δικτύων, με το δικό σου σχηματισμό συναπτικών βαρών και τις δικές σου κατανομές στο χώρο ενεργοποίησης» (Churchland, 1994, p.196).

Για τον Churchland, το παραπάνω γεγονός δεν μας οδηγεί αναπόφευκτα στην παραδοχή ότι υπάρχει κάτι μη φυσικό σε αυτές τις καταστάσεις, κάτι που υπερβαίνει τις αναπαραστάσεις των φυσικών επιστημών. Για να υποστηρίξει τη θέση του μας καλεί να εξετάσουμε τη λειτουργία του ιδιοδεκτικού συστήματος (proprioceptive system). Μέσω αυτού του δικτύου αξόνων διαθέτουμε πληροφοριακή πρόσβαση στο φυσικό σχηματισμό του σώματός μας και των μελών του. Εκατομμύρια αισθητήρες στους μυς μεταφέρουν σχετικές πληροφορίες στον εγκέφαλο. Έτσι, η ιδιάζουσα γνώση του σχηματισμού του σώματος μας και των κινήσεών του αποτελεί αποκλειστικά προσωπική εμπειρία, από τη στιγμή που μόνο ο εγκέφαλός μας έχει αυτές τις σχετικές αιτιακές συνδέσεις με το σώμα μας. Όπως και πριν οι άλλοι πρέπει να χρησιμοποιήσουν διαφορετικά μέσα – όραση, αφή κλπ - για να αποκτήσουν αυτή τη γνώση.

Αυτό το παράδειγμα ανοίγει το δρόμο για ένα πλήθος άλλων παραδειγμάτων. Η αιτιώδης σύνδεση που έχουμε με το στομάχι μας, με τους μυς στο δέρμα μας, με τα πνευμόνια μας, ο άμεσος τρόπος με τον οποίο αισθανόμαστε την καούρα, την ανατριχίλα, το κρύωμα, δεν μπορεί να γίνει κτήμα εξωτερικών παρατηρητών. Οι άλλοι μπορεί να διακρίνουν τις όρθιες τρίχες, τον ξερό βήχα και άλλες ενδεικτικές συμπεριφορές, αλλά ποτέ να αποκτήσουν την αδιαμεσολάβητη πρόσβαση που διαθέτει το κάθε άτομο προσωπικά σε αυτές τις εμπειρίες.

«Η ύπαρξη μιας προσωπικής, σε πρώτο πρόσωπο, επιστημολογικής πρόσβασης σε κάποιο φαινόμενο δε σημαίνει ότι το εν λόγω φαινόμενο είναι εγγενώς μη φυσικό», σημειώνει ο Churchland. Τα εμπλεκόμενα σωματικά φαινόμενα που αναφέρθηκαν είναι παραδειγματικώς φυσικά. Αυτός ο αυτο-εστιαζόμενος τρόπος γνώσης της ίδιας μας της

εσωτερικής κατάστασης αποτελεί τμήμα της ρύθμισης του σώματος κάθε πλάσματος και είναι απαραίτητος για την επιβίωσή του (Churchland, 1994, p.198).

Οι αυτο-συνδεδεμένοι τρόποι γνώσης έχουν ως γνωστικό αντικείμενο τα ίδια φυσικά πράγματα και γεγονότα που μπορούν να γίνουν γνωστά μέσω άλλων ετερο-συνδεδεμένων τρόπων. Η διαφορά έγκειται όχι στο περιεχόμενο της γνώσης, στο τι γνωρίζουμε, αλλά στον τρόπο με τον οποίο το γνωρίζουμε. Από την πλευρά του υποκειμένου μέσω των αυτο-συνδεδεμένων τρόπων (σωματοαισθητηριακό σύστημα), από την πλευρά του παρατηρητή μέσω ετερο-συνδεδεμένων τρόπων (οπτικό σύστημα).

Έτσι, και στο παράδειγμα του Nagel, ακόμα και να διαθέταμε τις απαραίτητες γνώσεις της νευροεπιστήμης σχετικά με τους εγκεφάλους των νυχτερίδων, δε θα μπορούσαμε να μοιραστούμε τις προσωπικές αισθητηριακές της εμπειρίες. Δεν έχουμε, για παράδειγμα, τις κατάλληλες αιτιώδεις συνδέσεις έτσι ώστε να εντοπίσουμε έντομα με τον ήχο, όπως κάνει μια νυχτερίδα.

Για τον Churchland, ο ισχυρισμός ότι οι νοητικές καταστάσεις έχουν μη φυσικές ιδιότητες - αν και παραδεκτός - δεν μπορεί να στηριχτεί στο επιχείρημα του Nagel. Σχετικά με την ύπαρξη μη φυσικών ιδιοτήτων, «κάποιο άλλο επιχείρημα πρέπει να επιφορτιστεί με την αποκάλυψή τους». Πάλι το πρόβλημα νου-σώματος έρχεται στο προσκήνιο. Ακόμα κι αν δεχτούμε την ύπαρξη μη φυσικών ιδιοτήτων πώς θα μπορούσαν αυτές να αλληλεπιδράσουν με φυσικά συμβάντα, όπως οι αυτοσυνδεδεμένες επιστημικές οδοί; Ο Churchland εντάσσει την εμφάνιση των νευρικών οδών σαν στάδιο της βιολογικής εξέλιξης που είχε σκοπό την ενοποίηση των εσωτερικών μας, κινητικών και αισθητηριακών, φυσιολογικών δραστηριοτήτων (Churchland, 1994, pp.199-200).

Κεφάλαιο 5 : Εναλλακτικές προσεγγίσεις

Εκτός από την αντιπαράθεση που εξετάστηκε, μεταξύ του Roger Penrose και του Paul Churchland, η παρούσα εργασία ασχολείται με κάποιες ακόμα σημαντικές συνεισφορές στη διαρκή συζήτηση για τη φύση της συνείδησης. Διάφοροι μελετητές και ερευνητές έχουν προτείνει εναλλακτικές θεωρίες και κριτικές προσεγγίσεις που εμπλουτίζουν το υπό εξέταση ερευνητικό πεδίο.

5.1 Η θεωρία του Daniel Dennett

Ο Daniel Dennett, διακεκριμένος φιλόσοφος και επιστήμονας, είναι γνωστός για τη συμβολή του στον τομέα της φιλοσοφίας του νου, με σημαντικό συγγραφικό έργο σχετικά με την *προθετικότητα* και το ρόλο της γλώσσας στη διαμόρφωση της συνείδησης. Μεταξύ άλλων, ο Dennett ασχολείται εκτενώς με το τεστ Turing και τη δυνατότητα μιας μηχανής να θεωρηθεί σκεπτόμενη ή να επιδειξει ευφυή συμπεριφορά. Ενώ ο Dennett αναγνωρίζει την αξία του τεστ Turing ως πειράματος σκέψης και πρακτικού εργαλείου για την αξιολόγηση της νοημοσύνης των μηχανών, επισημαίνει επίσης τους περιορισμούς του. Υποστηρίζει ότι η επιτυχία στο τεστ δεν συνεπάγεται απαραίτητα ότι μια μηχανή διαθέτει πραγματική συνείδηση ή κατανόηση. Αντίθετα, αποδεικνύει την ικανότητα της μηχανής να μιμείται ευφυή συμπεριφορά και να παράγει αληθοφανείς απαντήσεις σε ένα πλαίσιο συνομιλίας.

Στο άρθρο του «Can machines think?» (2004), ο Dennett εστιάζει στην περίπτωση του τεστ Turing, υπερασπιζόμενος την καθοριστική συμβολή του στην ανάδειξη των ορίων της τεχνητής νοημοσύνης. Στο έργο του αντικρούει διαφορετικές κριτικές που γίνονται στον Turing, θεωρώντας ότι όλες εστιάζουν σε λανθασμένες πτυχές του τεστ και υποτιμούν τη σημασία του.

Η διορατικότητα του τεστ έγκειται στον αμερόληπτο τρόπο που αυτό τίθεται. Αντί να λαμβάνονται υπόψη χαρακτηριστικά όπως το μαλακό δέρμα, το ζεστό αίμα, οι εκφράσεις του προσώπου κλπ, χαρακτηριστικά τα οποία δεν αποτελούν ουσιώδη δείγματα ευφυίας, αυτό που αξιολογείται είναι η ικανότητα για κατανόηση και ουσιαστική σκέψη πάνω σε απαιτητικά προβλήματα. Η ικανότητα, με άλλα λόγια, για διεξαγωγή μιας διεξοδικής και εκ βαθέων συζήτησης. Μια πιθανή επικράτηση στο τεστ Turing θα σήμαινε ότι η εκάστοτε υπολογιστική μηχανή θα ήταν σε θέση να επιδειξει οποιαδήποτε άλλη νοήμονα συμπεριφορά, ενώ μια πιθανή αποτυχία δεν σημαίνει και απουσία νοημοσύνης. Ίσως – παραφράζοντας τον Dennett - ο υπολογιστής να μην ήταν ικανός να εκτιμήσει την αισθητική απόλαυση που προσφέρει μια παράσταση μπαλέτου, να κατανοήσει τις τελευταίες εξελίξεις στην κβαντομηχανική θεωρία ή να προσφέρει μια οριστική λύση για την πυρηνική απειλή, θα μπορούσε να θεωρηθεί, ωστόσο, μια ευφυή, σκεπτόμενη οντότητα (Dennett, 2004).

Ο Dennett υποστηρίζει ότι ένας τέτοιος «επιτυχημένος» διαλογικός υπολογιστής πρέπει να κατέχει πληθώρα γνώσης για τον κόσμο (world knowledge). Προς επίρρωση του επιχειρήματός του δανείζεται το παράδειγμα των δυο παρακάτω προτάσεων, όπως τέθηκε από έναν επιστήμονα TN, τον Terry Winograd.

1. Η επιτροπή αρνήθηκε στην ομάδα την άδεια παρέλασης επειδή προωθούσαν τη βία.
2. Η επιτροπή αρνήθηκε στην ομάδα την άδεια παρέλασης επειδή φοβόντουσαν τη βία.

Η σύγκυση που μπορεί να προκληθεί οφείλεται στη διττή σημασία που μπορεί να αποδοθεί στην ανωνυμία που υπονοούν τα ρήματα «προωθούσαν» και «φοβόντουσαν». Ενώ και οι δυο αναγνώσεις είναι γραμματικά «νόμιμες», μόνο η μία βγάζει όντως νόημα στον κόσμο που ζούμε, ότι η επιτροπή φοβάται τη χρήση βίας από την ομάδα (Dennett, 2004, p.298).

Όταν τέτοιου είδους προτάσεις εισάγονται στη συζήτηση με έναν υπολογιστή, οι απλοί κανόνες γραμματικής και λεξιλογίου δεν αρκούν, ώστε αυτός να προσφέρει μια ευφυή απάντηση. Η σωστή ανάγνωση εξαρτάται, εν τέλει, από τη γενικότερη αντίληψη για τον κόσμο, τις κοινωνικές και πολιτικές συνθήκες, τις διαδικασίες που ακολουθούν οι επιτροπές και μια σειρά συναφών γνώσεων.

Για τον Dennett δε διαφαίνεται κάποια πιθανότητα τα επόμενα χρόνια για έναν υπολογιστή να επιτύχει στο τεστ, «ακόμα κι αν κερδίσει το Παγκόσμιο Πρωτάθλημα Σκακιού ή το Νόμπελ Φυσικής». Ωστόσο, δε θεωρεί αδύνατη κατ'αρχήν, μια ενδεχόμενη επιτυχία. Σημειώνει ότι υπάρχουν σημαντικές μορφές σκέψης ή νοημοσύνης, που δεν είναι καλά προετοιμασμένες να παίξουν και να κερδίσουν το παιχνίδι της μίμησης. Το γεγονός ότι η επικράτηση στο τεστ φαίνεται σχεδόν απίθανη δε σημαίνει ότι δεν υπάρχουν μηχανές που ήδη παρουσιάζουν κάποια από τα σημαντικά χαρακτηριστικά της ενσυνείδητης σκέψης. Για αυτές τις μηχανές το αρχικό ερώτημα «μπορούν να σκέφτονται;» μοιάζει ανώφελο, καθώς η απάντηση βρίσκεται στην εσωτερική δομή τους. «Μόνο μια λεπτομερής ματιά στο τι κάνουν και πως είναι δομημένες θα αποκαλύψει το ενδιαφέρον τους», και το Turing τεστ – ως ένα μη επιστημονικό τεστ - δε φαίνεται ικανό να το αποκαλύψει (Dennett, 2004, p.309).

Ο Dennett υποστηρίζει ότι το Turing Test εστιάζει κυρίως στην εξωτερική συμπεριφορά της μηχανής, παραμελώντας σημαντικές εσωτερικές διεργασίες που συμβάλλουν στην ανθρώπινη νοημοσύνη, όπως η υποκειμενική εμπειρία, τα συναισθήματα και η αυτογνωσία. Υποστηρίζει ότι αυτές οι εσωτερικές πτυχές αποτελούν βασικά συστατικά της ανθρώπινης σκέψης, τα οποία οι μηχανές δεν έχουν ακόμη καταφέρει να αναπαράγουν πειστικά.

Σε ένα υστερόγραφο του άρθρου του «Can machines think», υποστηρίζει ότι οι άνθρωποι θα μπορούσαν να προγραμματίσουν έναν υπολογιστή με αυτογνωσία, αλλά υπό μια στενή και απλουστευτική έννοια της λέξης. Με αυτήν την έννοια και ένας αστακός διαθέτει την ικανότητα της αυτογνωσίας, εφόσον «όταν πεινάσει, τρώει κάτι, αλλά ποτέ τον εαυτό του». Μπορεί δηλαδή να προβεί σε διάκριση μεταξύ του εαυτού του και του υπόλοιπου κόσμου. Αντίστοιχα, ένας υπολογιστής αυτοελέγχεται και αυτοεπιτηρείται.

Η καθημερινή όμως χρήση της έννοιας της αυτογνωσίας έχει συνήθως να κάνει «με αυτό το ιδιαίτερο εσωτερικό φως, με τον προσωπικό δρόμο του καθενός, που κανένας άλλος δεν μπορεί να μοιραστεί και που παραμένει διαρκώς έξω από τα όρια της επιστήμης των υπολογιστών», σημειώνει με γλαφυρότητα ο Dennett. Για τον Dennett κάτι τέτοιο αποτελεί

ψευδαίσθηση της κοινής λογικής, ψευδαίσθηση που προσπαθεί να εξαλείψει με την χρήση παραδειγμάτων που απελευθερώνουν τη φαντασία (intuition pumps), όπως το παρακάτω.

Στο συγκεκριμένο παράδειγμα μας καλεί να φέρουμε στο μυαλό μας την εικόνα ενός υπολογιστή, προσωπικού ή μεγάλων διαστάσεων, και να εξετάσουμε το εσωτερικό του. Τίποτα από το υλισμικό του και τα διαφορετικά συστήματα κυκλωμάτων δε μας οδηγεί στο να συμπεράνουμε ότι μπορεί μια τέτοια μηχανή να διαθέτει συνείδηση. Με παρόμοιο τρόπο, η αφαίρεση του κρανίου και η εξέταση του εγκεφάλου και των εσωτερικών του περιοχών δεν μπορεί να μας πείσει για το ότι η συνείδηση εδράζεται εντός του. Σε κανένα επίπεδο αυτοψίας ένας εγκέφαλος ή ένας υπολογιστής δεν μοιάζει με την έδρα της συνείδησης. «Αν θέλετε να συλλάβετε πώς ένας υπολογιστής θα μπορούσε να έχει συνείδηση, δεν είναι τελικά πιο δύσκολο από το να κατανοήσετε πώς ένας εγκέφαλος θα μπορούσε να έχει συνείδηση», μας πληροφορεί ο Dennett (Dennett, *Intuition pumps and other tools for thinking*, 2013).

Βαρύνουσα θέση στη σκέψη του Dennett κατέχει, επίσης, η έννοια σκοπιμότητας στην ανθρώπινη σκέψη. Η *σκοπιμότητα* (ή προθετικότητα ή αναφορικότητα) αναφέρεται στη λειτουργία των νοητικών καταστάσεων, κατά την οποία οι πεποιθήσεις, οι επιθυμίες και οι αντιλήψεις μας κατευθύνονται, αναφέρονται προς συγκεκριμένα αντικείμενα ή έννοιες. Για παράδειγμα, όταν έχουμε μια πεποίθηση για κάτι ή επιθυμούμε κάτι, αυτές οι νοητικές καταστάσεις είναι σκόπιμες καθώς αφορούν συγκεκριμένα πράγματα. Η ανθρώπινη σκέψη είναι εγγενώς σκόπιμη, με την έννοια ότι οι πεποιθήσεις, οι επιθυμίες, οι αντιλήψεις και άλλες νοητικές μας καταστάσεις έχουν νόημα επειδή κατευθύνονται προς κάτι στον κόσμο. Έχουμε μια υποκειμενική κατανόηση των αντικειμένων ή των εννοιών που αφορούν οι σκέψεις μας. Αυτή η υποκειμενική εμπλοκή με τον κόσμο προσδίδει στις σκέψεις και τις νοητικές μας καταστάσεις ένα βαθύτερο επίπεδο σημασίας. Ενώ οι μηχανές μπορούν να επεξεργάζονται πληροφορίες και να παράγουν αποτελέσματα, στερούνται γνήσιας σκοπιμότητας. Μπορούν να χειρίζονται σύμβολα και να εκτελούν εργασίες βάσει προγραμματισμένων κανόνων, αλλά δεν διαθέτουν εγγενή κατανόηση ή υποκειμενική εμπλοκή με τον κόσμο.

Για αυτόν, η νοημοσύνη δεν είναι μια μοναδική, ενοποιημένη οντότητα, αλλά μάλλον μια συλλογή από διάφορες γνωστικές ικανότητες ή "πολλαπλές νοημοσύνες". Υποστηρίζει ότι οι διάφορες πτυχές της ανθρώπινης νοημοσύνης, όπως η μαθηματική σκέψη, η γλωσσική κατανόηση, η χωρική αντίληψη, η συναισθηματική νοημοσύνη και η κοινωνική κατανόηση μπορούν να ποικίλλουν. Οι μηχανές μπορεί να υπερέχουν σε συγκεκριμένες γνωστικές λειτουργίες και εργασίες, όπως οι πολύπλοκοι υπολογισμοί ή η γλωσσική επεξεργασία, ξεπερνώντας τις ανθρώπινες επιδόσεις σε αυτούς τους τομείς. Ωστόσο, υποστηρίζει ότι οι μηχανές εξακολουθούν να στερούνται της ευρύτερης κατανόησης και συνείδησης που διαθέτουν οι άνθρωποι.

Ο Dennett προχωρεί σε διάκριση μεταξύ αυτού που ονομάζει *ικανότητα* και *κατανόηση*. Η ικανότητα αναφέρεται στην ικανότητα εκτέλεσης συγκεκριμένων διεργασιών ή αλγορίθμων, ενώ η κατανόηση περιλαμβάνει μια βαθύτερη κατανόηση και επίγνωση του νοήματος και των συνεπειών αυτών των διεργασιών. Οι μηχανές μπορεί να παρουσιάζουν

υψηλά επίπεδα ικανότητας σε συγκεκριμένους τομείς, αλλά δεν διαθέτουν τη συνολική κατανόηση και το πλαίσιο που οι άνθρωποι προσδίδουν στις ενέργειές τους.

5.2 Η Ολοκληρωμένη Θεωρία Πληροφοριών

Η «Ολοκληρωμένη Θεωρία Πληροφοριών» (Integrated Information Theory - ΙΙΤ) που προτάθηκε αρχικά από τον Giulio Tononi το 2004, παρέχει ένα διαφορετικό θεωρητικό πλαίσιο για την κατανόηση της συνείδησης. Από το 2004 μέχρι σήμερα, ο ίδιος, μαζί με τον Koch, δημοσίευσαν πολλές εκδοχές της ΙΙΤ, με μικρές αλλαγές ή βελτιώσεις μεταξύ τους. Η γενική ιδέα παραμένει αμετάβλητη. Υπάρχουν πολλά ερωτήματα σχετικά με τη συνείδηση που φαίνεται να είναι πολύ δύσκολο να απαντηθούν. Πώς μπορεί η ύλη να δημιουργήσει συνείδηση; Γιατί αισθανόμαστε μια «ερυθρότητα» όταν βλέπουμε ένα όμορφο λουλούδι και όχι κάτι άλλο; Γιατί ορισμένες δομές του εγκεφάλου συνδέονται με την εμπειρία, αλλά άλλες, όπως η παρεγκεφαλίδα, όχι; Ποια ζώα έχουν συνείδηση; Μπορούν οι υπολογιστές να έχουν συνείδηση; Αν δεν θέλουμε οι απαντήσεις να βασίζονται στη θρησκεία, την παράδοση ή τη διαίσθηση, τότε θα πρέπει να δώσουμε μια επιστημονική εξήγηση της συνείδησης. Ωστόσο, υπάρχουν πολλές δυσκολίες όταν προσπαθούμε να περιγράψουμε την υποκειμενική εμπειρία με επιστημονικούς τρόπους, όπως αναφέρθηκε στο προηγούμενο μέρος της παρούσας εργασίας. Οι Tononi και Koch πρότειναν μια νέα μέθοδο αντιμετώπισης του προβλήματος. Αντί να ξεκινούν από τον εγκέφαλο και να προσπαθούν να βρουν πώς αυτός γεννά την εμπειρία, όπως κάνουν οι περισσότεροι επιστήμονες, ξεκίνησαν από την ίδια τη συνείδηση, από τις βασικές ιδιότητες της φαινομενικής εμπειρίας, και στη συνέχεια προσπάθησαν να βρουν τι είδους φυσικοί μηχανισμοί είναι απαραίτητοι για να τις έχει ένα σύστημα.

Η εκκίνηση από τη φαινομενολογία και όχι από τους μηχανισμούς του εγκεφάλου είναι μια ελπιδοφόρα προσπάθεια. Εξάλλου, όταν μιλάμε για νοημοσύνη, πρέπει να λαμβάνουμε υπόψη την υποκειμενική φύση των νοητικών φαινομένων.

Οι Tononi και Koch καταλήγουν στο συμπέρασμα ότι υπάρχουν 5 αξιώματα, δηλαδή 5 βασικές, αυταπόδεικτες ιδιότητες της εμπειρίας. Πριν καταλήξουν ποια είναι αυτά τα 5 αξιώματα, παραθέτουν ορισμένες ικανότητες που συνήθως συνδέουμε με τη συνείδηση και που οι ερευνητές της TN προσπαθούν να μιμηθούν στις μηχανές, αλλά δεν είναι απαραίτητες γι' αυτήν. Στο «Can Machines Be Conscious?» (2008), αναφέρουν, τι θεωρούν περιττό για τη συνείδηση :

1) Αισθητηριακή εισροή και κινητική εκροή.

Όταν ονειρευόμαστε, παρόλο που είμαστε αποσυνδεδεμένοι από το περιβάλλον και οι μύες μας είναι σε μεγάλο βαθμό παράλυτοι, έχουμε συνείδηση πολλών από τα όνειρά μας. Οι τυφλοί άνθρωποι είναι σε θέση να φαντάζονται και να ονειρεύονται εικόνες και οι ασθενείς με σύνδρομο “locked-in”, αν και πλήρως παράλυτοι, είναι εξίσου συνειδητοποιημένοι με τα υγιή άτομα. Επομένως, η αλληλεπίδραση με το περιβάλλον δεν είναι απαραίτητη για τη συνείδηση.

2) Συναισθήματα

Οι ασθενείς με βλάβες στον προμετωπιαίο φλοιό παρουσιάζουν μια επίπεδη, χωρίς συναισθήματα επιρροή. Παρόλα αυτά, βιώνουν ένα τραγούδι ή ένα τοπίο, με τον ίδιο τρόπο που τα βιώνουμε κι εμείς.

3) Επιλεκτική προσοχή

Ένα άτομο μπορεί να βιώσει συνειδητά ένα γεγονός, χωρίς να του δώσει προσοχή. «Όταν είστε συγκεντρωμένοι σε μια καθηλωτική ταινία, το περιβάλλον σας δεν περιορίζεται σε ένα τούνελ». Επιπλέον, όπως έχουν δείξει πολλά πειράματα, μπορεί κανείς να παρακολουθεί ένα αντικείμενο (να το επεξεργάζεται κατά προτίμηση) χωρίς να το αντιλαμβάνεται συνειδητά.

4) Μνήμη

Ένας ασθενής με επιληπτικές κρίσεις ο οποίος υποβλήθηκε σε αμφίπλευρη μεσοκροταφική λοβεκτομή, δεν μπορούσε να σχηματίσει νέες μακροπρόθεσμες μνήμες. Αυτό δεν τον μετέτρεψε σε ζόμπι. Πέθανε πριν από λίγα χρόνια, ικανός να έχει συνειδητές εμπειρίες, σε όλη του τη ζωή.

5) Αυτο-ανασκόπηση

«Όταν απορροφούμαστε σε κάποια έντονη αντιληπτική εργασία -όπως το να παίζουμε ένα γρήγορο βιντεοπαιχνίδι, να στρίβουμε με μια μοτοσυκλέτα μέσα στην κινούμενη κυκλοφορία ή να τρέχουμε κατά μήκος ενός ορεινού μονοπατιού- έχουμε έντονη συνείδηση του εξωτερικού κόσμου, χωρίς καμία ανάγκη για προβληματισμό ή ενδοσκόπηση». Οι μελέτες νευροαπεικόνισης δείχνουν ότι οι βλάβες στον προμετωπιαίο, που εμπλέκεται στην αυτοαντιπροσώπευση, παρόλο που προκαλούν σοβαρές γνωστικές ελλείψεις, δεν μειώνουν την ικανότητα των ασθενών να έχουν συνειδητές εμπειρίες.

6) Γλώσσα

Πιστεύουμε ότι οι άλλοι άνθρωποι έχουν συνείδηση, επειδή τους ακούμε να περιγράφουν τις εμπειρίες τους. Όμως η ομιλία δεν είναι απαραίτητη για τη συνείδηση. Οι ασθενείς με βλάβες στην περιοχή Broca δεν χάνουν την ικανότητά τους για υποκειμενικές εμπειρίες. Τα βρέφη και οι πίθηκοι δεν μπορούν να μιλήσουν, αλλά μπορούν να εκφράσουν ότι έχουν συνείδηση με άλλους τρόπους.

Τι είναι λοιπόν απαραίτητο; Ποιες είναι οι βασικές ιδιότητες της συνείδησης; Οι Tononi και Koch πρότειναν 5 αξιώματα που αντιπροσωπεύουν κάθε πιθανή εμπειρία και όπως είπαν: «Τα αξιώματα της ΔΙΕ είναι υποθέσεις σχετικά με τη δική μας εμπειρία που αποτελούν το σημείο εκκίνησης της θεωρίας. Ιδανικά, τα αξιώματα είναι ουσιώδη (ισχύουν για όλες τις εμπειρίες), πλήρη (περιλαμβάνουν όλες τις βασικές ιδιότητες που μοιράζεται κάθε εμπειρία), συνεπή (δεν έχουν αντιφάσεις) και ανεξάρτητα (δεν μπορούν να συναχθούν μεταξύ τους)».

1) Το πρώτο είναι το αξίωμα της *ύπαρξης*. Το ότι έχουμε συνείδηση είναι το μόνο πράγμα για το οποίο είμαστε απολύτως βέβαιοι. Γράφω αυτό το δοκίμιο αυτή τη στιγμή και αυτό είναι αναμφισβήτητα αληθινό για μένα. Η εμπειρία μου υπάρχει και υπάρχει για μένα,

από τη δική μου εγγενή προοπτική, ανεξάρτητα από το τι σκέφτεται ένας εξωτερικός παρατηρητής.

2) Το δεύτερο είναι το αξίωμα της *σύνθεσης*. Η συνείδηση είναι δομημένη. Η εμπειρία μας αποτελείται από πολλές φαινομενολογικές διακρίσεις, οι οποίες επίσης υπάρχουν. Μέσα στην ίδια εμπειρία, μπορώ να δω τον υπολογιστή μου, τις μαύρες λέξεις που γράφω σε λευκό φόντο, το μπλε μπουκάλι νερό δίπλα μου και το κρεβάτι μου στο δωμάτιό μου, ενώ ακούω τον αγαπημένο μου ραδιοφωνικό σταθμό.

3) Το τρίτο είναι το αξίωμα της *πληροφόρησης*. Κάθε εμπειρία αποτελείται από ένα συγκεκριμένο σύνολο φαινομενικών διακρίσεων, οι οποίες την καθιστούν αυτό που είναι και όχι κάτι άλλο. Η εμπειρία μου εδώ και τώρα είναι συγκεκριμένη και επειδή είναι έτσι όπως είναι, διαφέρει από άλλες εμπειρίες. Για τον Tononi, η πληροφορία είναι η μείωση της αβεβαιότητας όταν επιλέγεται κάτι, ανάμεσα σε άλλες πιθανές επιλογές. «Όταν βλέπετε την οθόνη να σκοτεινιάζει, εισέρχεστε σε μία από έναν τεράστιο αριθμό καταστάσεων: το να βλέπετε μια σκοτεινή οθόνη σημαίνει ότι δεν βλέπετε μια μπλε, κόκκινη ή πράσινη οθόνη, το Άγαλμα της Ελευθερίας, μια εικόνα από το ρεσιτάλ πιάνου του παιδιού σας ή οποιοδήποτε άλλο από τα αμέτρητα πράγματα που έχετε δει ποτέ ή θα μπορούσατε ποτέ να δείτε. Για εσάς, "σκοτάδι" δεν σημαίνει απλώς το αντίθετο του φωτός, αλλά επίσης, και ταυτόχρονα, κάτι διαφορετικό από τα χρώματα, τα σχήματα, τις μυρωδιές ή οποιοδήποτε μείγμα των παραπάνω» (Koch & Tononi, 2008).

4) Το τέταρτο είναι το αξίωμα της *ολοκλήρωσης*. Η συνείδηση είναι ενιαία. Δεν μπορούμε να αναγάγουμε τις εμπειρίες μας σε μη αλληλεξαρτώμενα υποσύνολα φαινομενικών διακρίσεων. Τα χρώματα και τα σχήματα που δομούν την οπτική μου εμπειρία βιώνονται μαζί. Δεν μπορώ να βιώσω το σχήμα του μπουκαλιού με το νερό δίπλα μου, ξεχωριστά από το μπλε χρώμα του. Οποιαδήποτε πληροφορία έχω συνείδηση παρουσιάζεται εξ ολοκλήρου και ολοκληρωτικά στο μυαλό μου. Δεν μπορεί να υποδιαριεθεί. (Koch, 2009)

5) Τέλος, η συνείδηση έχει την ιδιότητα του αποκλεισμού. Ακριβώς όπως η συνείδηση προσδιορίζει ορισμένα πράγματα, αποκλείει άλλα. Βιώνω τη μουσική από το ραδιόφωνο, αλλά δεν βιώνω την αρτηριακή μου πίεση ενώ ακούω το τραγούδι ούτε βιώνω το μπουκάλι με το νερό ασπρόμαυρο, χωρίς το μπλε χρώμα του. Κάθε εμπειρία έχει το σύνολο των φαινομενικών διακρίσεων που έχει, ούτε λιγότερο ούτε περισσότερο και ρέει με την ταχύτητα που ρέει, ούτε πιο γρήγορα ούτε πιο αργά.

Παρατηρούμε από όσα αναφέρθηκαν ότι πολλές πτυχές της ΠΤ θα έβρισκαν σύμφωνο τον Paul Churchland στην ανάλυσή του για τις διαστάσεις της συνείδησης.

Κεφάλαιο 6 : Με το βλέμμα στο μέλλον

Η συνείδηση φαίνεται να αποτελεί το κυριότερο φιλοσοφικό ζήτημα του 21^{ου} αιώνα. Κάτι τέτοιο επιβεβαιώνεται από τη δημοσίευση άρθρων σε επιστημονικά περιοδικά, από την προθυμία ιδρυμάτων να στεγάσουν διεπιστημονικά ερευνητικά κέντρα επιχορηγώντας αντίστοιχα συνέδρια, αλλά και από την τάση των εκδοτικών οίκων για κυκλοφορία σχετικών βιβλίων στο εμπόριο.

Παρομοιάζοντας τη φιλοσοφική διαμάχη με αγώνα πυγμαχίας, ο διακεκριμένος φιλόσοφος Thomas Nagel, αναφέρει χαρακτηριστικά ότι «οι φιλόσοφοι μοιάζουμε με γηράσκοντες επαγγελματίες πυγμάχους που μονίμως ανταλλάσσουν τα ίδια χτυπήματα» (Μαραγκός, 1995). Στον αγώνα όμως με έπαθλο την εξήγηση της συνείδησης καμία άποψη δεν έχει τεθεί οριστικά εκτός μάχης.

Κάποιες θεωρίες έχουν χάσει, βέβαια, την επεξηγηματική τους ισχύ λόγω της προόδου της σύγχρονης επιστήμης. Ελάχιστοι, πλέον, οι υποστηρικτές του καρτεσιανού οντολογικού δυισμού σώματος-νου, όπως και οι οπαδοί του μεθοδολογικού συμπεριφορισμού. Η ανθρώπινη συμπεριφορά δεν είναι απλώς πραγμάτωση ανακλαστικών διαθέσεων, αλλά διέπεται και από εσωτερικά αίτια. Σε ό,τι αφορά τις διανοητικές καταστάσεις και λειτουργίες ως αποτελέσματα εξωτερικών ερεθισμάτων και ως αίτια της συμπεριφοράς, ο μονισμός φαίνεται να είναι η επικρατούσα άποψη στο επίπεδο της οντολογίας, άποψη που ανάγει την ύλη στη μοναδική ουσία του σύμπαντος. Οι διανοητικές καταστάσεις και λειτουργίες είναι έκφανση ή σύστοιχο καταστάσεων και λειτουργιών του Νευρικού Συστήματος.

6.1 Σύνοψη των ευρημάτων

Μετά από μια συνολική εξέταση της αντιπαράθεσης μεταξύ του Roger Penrose και του Paul Churchland σχετικά με τη φύση της συνείδησης, καθώς και των θεωριών και των επιχειρημάτων που διατύπωσαν στα αντίστοιχα έργα τους, "Σκιές του Νου" και "Η Μηχανή της Λογικής, η Θέση της Ψυχής", μπορούν να εξαχθούν μερικά βασικά συμπεράσματα.

Ο Penrose υποστηρίζει ότι ο εγκέφαλος δεν μπορεί να προσεγγιστεί ως ένας πολύπλοκος υπολογιστής διότι οι λογικές διεργασίες που εκτελεί δεν είναι αλγοριθμικές. Βασιζόμενος στα θεωρήματα μη πληρότητας του Godel, οδηγείται στο συμπέρασμα ότι οι υπολογιστές δεν μπορούν να αποκτήσουν συνείδηση. Μια καλά τεκμηριωμένη θεωρία για τη συνείδηση δεν μπορεί να αρκείται στα κλασικά υπολογιστικά μοντέλα και η κατανόηση των κβαντικών φαινομένων κρίνεται απαραίτητη για την εξήγηση του υποκειμενικού της στοιχείου. Η Ενορχηστρωμένη Θεωρία της Συνείδησης του Penrose εστιάζει στις κβαντικές διεργασίες μέσα στους μικροσωληνίσκους του εγκεφάλου, οι οποίοι συμβάλλουν στην εμφάνιση της συνειδητής εμπειρίας. Ωστόσο, η θεωρία του έχει αντιμετωπίσει επικρίσεις σχετικά με τη δυνατότητα εμπειρικού ελέγχου των προτάσεών της, αλλά και σχετικά με τη δυνατότητα διατήρησης της κβαντικής συνοχής σε μακροσκοπικό επίπεδο.

Από την άλλη πλευρά, η επιστημονική σκοπιά του Churchland δίνει έμφαση στο ρόλο των νευροϋπολογιστικών αναλύσεων και των δικτύων επανατροφοδότησης στην κατανόηση του φαινομένου της συνείδησης. Διατηρώντας μια πιο αναγωγιστική στάση, προτείνει ότι η εξέταση των νευρωνικών μηχανισμών που διέπουν τις γνωστικές διεργασίες και της δυνατότητας του εγκεφάλου για μαζικά παράλληλη κατανεμημένη επεξεργασία, μπορεί να αποβεί καθοριστικής σημασίας για την εξήγηση της ενσυνείδητης αντίληψης. Η εστίαση αυτή του Churchland στα νευρωνικά δίκτυα έχει ασκήσει επιρροή στη γεφύρωση του χάσματος μεταξύ νευροεπιστήμης και φιλοσοφίας του νου. Μια σημαντική κριτική στη θεωρία του, μεταξύ άλλων, αφορά τις υποκειμενικές πτυχές της συνείδησης (qualia) που ενδέχεται να μην μπορούν να συλληφθούν πλήρως από αμιγώς φυσικές ή υπολογιστικές εξηγήσεις.

6.1.1 Κοινοί τόποι Penrose – Churchland

Από την εξέταση των θεωριών των δυο συγγραφέων προκύπτουν μερικά σημεία σύγκλισης μεταξύ τους. Πρώτον, τόσο ο Penrose όσο και ο Churchland αναγνωρίζουν την πολυπλοκότητα και τη σημασία του προβλήματος νους-σώμα. Διαβλέπουν ότι η κατανόηση του τρόπου με τον οποίο οι υποκειμενικές εμπειρίες και η συνειδητή επίγνωση προκύπτουν από φυσικές διεργασίες στον εγκέφαλο αποτελεί μια βαθιά πρόκληση. Συμφωνούν ότι κάθε ολοκληρωμένη θεωρία της συνείδησης πρέπει να αντιμετωπίσει αυτό το θεμελιώδες πρόβλημα. Αμφότεροι αναγνωρίζουν τη σημασία της επιστημονικής έρευνας, της παρουσίασης εμπειρικών στοιχείων και της διεξαγωγής πειραμάτων για την κατανόηση της νόησης. Ενώ οι προσεγγίσεις και οι ερμηνείες τους μπορεί να διαφέρουν, μοιράζονται τη δέσμευση για την επιστημονική διερεύνηση της συνείδησης.

Επιπλέον, και οι δύο φιλόσοφοι ενθαρρύνουν τις διεπιστημονικές προσεγγίσεις στη μελέτη της συνείδησης. Αναγνωρίζουν ότι η κατανόηση της συνείδησης απαιτεί συνεργασία διαφόρων γνωστικών πεδίων, όπως της φιλοσοφίας, της νευροεπιστήμης, των γνωσιακών επιστημών και της φυσικής.

Τόσο ο Penrose όσο και ο Churchland δίνουν έμφαση στο ρόλο που διαδραματίζει η φιλοσοφική έρευνα στην εξήγηση της φύσης της συνείδησης, καθώς οι φιλοσοφικές έννοιες και τα φιλοσοφικά πλαίσια, στα θεμέλια των οποίων στηρίζονται, συμπυκνώνουν αιώνες επιστημονικών ερευνών και θεωρητικού στοχασμού. Αν και μπορεί να διαφέρουν στις φιλοσοφικές τους θέσεις, μοιράζονται από κοινού την εκτίμηση για τον ρόλο της φιλοσοφίας στην εμβάθυνση της κριτικής σκέψης και στην ανάπτυξη χρήσιμων εννοιών και εξηγητικών εργαλείων.

Τούτων λεχθέντων, ένα συμπέρασμα που προκύπτει είναι ότι και οι δυο επιστήμονες υιοθετούν την προοπτική της ριζικής εμμένειας στις θεωρίες που αναπτύσσουν στα έργα τους. Τίποτα έξω από αυτόν τον κόσμο, τη σκέψη και τη γλώσσα δεν μπορεί να προσφέρει λύσεις στη συζήτηση για την ανάδυση της ενσυνείδητης αντίληψης. Η εναπόθεση των προσδοκιών μας σε μια έξωθεν δεσπόζουσα θέση για μια ολοκληρωμένη εξήγηση του φαινομένου, μόνο παραπλανητική μπορεί να αποβεί. Η διάθεση του Churchland να εμμένει στο νευροϋπολογιστικό μονοπάτι ως εννοιολογικό εργαλείο και η διάθεση του Penrose για

εξέταση των κβαντικών φαινομένων εντός ενός φυσικού συστήματος, του εγκεφάλου, καταδεικνύουν την ανάγκη για αναζήτηση λύσεων εντός του επιστημονικού πλαισίου. Παρά τις επιμέρους διαφορές τους, και οι δυο αποφεύγουν την αναγωγή των συνειδησιακών λειτουργιών σε κάτι πέραν αυτού του επιπέδου πραγματικότητας, πέραν του πεδίου όσων συμβαίνουν στον κόσμο. Η προσπάθεια να προσεγγίσουν τη συνείδηση με διαφορετικό τρόπο, σαν ένα φαινόμενο εκτός του κόσμου, θα συνιστούσε α-νοησία, καθώς θα προσέκρουε στα όρια της επιστημονικής σκέψης. Μια τέτοια απόπειρα επεξήγησης της ενσυνείδητης σκέψης με μεταφυσικούς όρους θα οδηγούσε, αναπόφευκτα, σε μια θρησκευτική, μυστικιστική, ή ηθική αντίληψη περί συνείδησης. Οι επιστημονικές έννοιες όπως και οι λέξεις, εκφράζουν απλώς γεγονότα, δεν μπορούν να εκφράσουν την ηθική, η οποία υπερβαίνει τις λέξεις και τον φυσικό κόσμο.

Αν και μια ολοκληρωμένη εξήγηση της συνείδησης και των σχετικών εκδηλώσεών της φαίνεται να μην βρίσκεται εντός των δυνατοτήτων της σύγχρονης επιστημονικής κατανόησης, δεν πρέπει να θεωρηθεί ότι θα βρίσκεται για πάντα εκτός της σφαίρας της επιστήμης. Όπως σημειώνει και ο Penrose στην εισαγωγή του βιβλίου του, «αν η επιστήμη είναι ακόμη ανίκανη να πει πολλά σημαντικά πράγματα σχετικά με τα ζητήματα του νου, τότε τελικά η επιστήμη πρέπει να διευρύνει το πεδίο εφαρμογής της ώστε να χωρέσει τέτοια ζητήματα, και ίσως ακόμη και να τροποποιήσει τις ίδιες τις διαδικασίες της» (Penrose, 1996, p.12). Στο συγκεκριμένο απόσπασμα, ο Penrose αναφέρεται, ουσιαστικά, στην αναγκαία - με τους όρους του Μπαλτά - *γραμματική ρήξη*, τη διεύρυνση δηλαδή του γραμματικού χώρου και την εννοιολογική επέκταση που απαιτούνται, έτσι ώστε οι νέες επιστημονικές θεωρίες και φιλοσοφικές θέσεις να εγκολπώσουν και να εγκολπωθούν στις σύγχρονες τεχνολογικές εξελίξεις. Η δουλειά της φιλοσοφίας του νου, αλλά και της φιλοσοφίας γενικότερα, παραμένει στον πυρήνα της η ίδια, όπως διατυπώθηκε από τον L.Wittgenstein στο περίφημο *Tractatus*: «η φιλοσοφία πρέπει να αποσαφηνίζει και να οριοθετεί αυστηρά τις σκέψεις που συνήθως είναι θολές και συγκεχυμένες».

6.1.2 Διαχωριστικές γραμμές Penrose – Churchland

Ένα σημαντικό σημείο διαφωνίας μεταξύ των δυο έχει να κάνει με τη φύση της ίδιας της συνείδησης. Ο Penrose υιοθετεί μια μη υπολογιστική άποψη, υποδηλώνοντας ότι η συνείδηση μπορεί να περιλαμβάνει μη υπολογίσιμες διαδικασίες και κβαντικά φαινόμενα που είναι πέρα από το πεδίο εφαρμογής της κλασικής φυσικής. Αντίθετα, ο Churchland δίνει έμφαση σε μια νευροϋπολογιστική προσέγγιση της συνείδησης, θεωρώντας την ως μια αναδυόμενη ιδιότητα της παράλληλης επεξεργασίας πληροφοριών στον εγκέφαλο. Αυτή η θεμελιώδης απόκλιση στις απόψεις τους για τη φύση της συνείδησης οδηγεί σε επακόλουθες διαφορές στις θεωρίες και τις εξηγήσεις τους. Σε σχέση με την ικανότητα ενσυνείδητης αντίληψης στις μηχανές τεχνητής νοημοσύνης, ο Penrose εκφράζει σκεπτικισμό για τη δυνατότητα δημιουργίας πραγματικά συνειδητών συστημάτων TN, υποστηρίζοντας ότι η συνείδηση περιλαμβάνει μη αλγοριθμικές διαδικασίες που δεν μπορούν να αναπαραχθούν από μηχανές. Αντίθετα, ο Churchland αναγνωρίζει τη δυνατότητα της TN να εμφανίζει μορφές συνείδησης μέσα από διαδικασίες που περιγράφηκαν λεπτομερώς στο 4^ο κεφάλαιο.

Μια ακόμα διαφορά προκύπτει από τις απόψεις τους για τον ρόλο της κβαντικής φυσικής στην εξήγηση της συνείδησης. Ο Penrose προτείνει μια θεωρία της κβαντικής συνείδησης, προτείνοντας ότι οι κβαντικές διεργασίες εντός των νευρικών δομών συμβάλλουν στην εμφάνιση της συνείδησης. Ο Churchland, από την άλλη πλευρά, τείνει να είναι επιφυλακτικός ως προς την άμεση εμπλοκή των κβαντικών φαινομένων στη συνείδηση και εστιάζει περισσότερο σε νευρωνικές και υπολογιστικές εξηγήσεις.

Ο Penrose και ο Churchland εκκινούν από διαφορετικές αφετηρίες για την ανάπτυξη της συλλογιστικής τους στη μελέτη του φαινομένου. Ο Penrose βασίζεται σε μαθηματικές και φυσικές αρχές, συμπεριλαμβανομένου του θεωρήματος της μη πληρότητας του Gödel, για να υποστηρίξει τα επιχειρήματα και τις θεωρίες του. Ο Churchland, από την άλλη πλευρά, δίνει μεγαλύτερη έμφαση στα εμπειρικά στοιχεία, στα μοντέλα νευρωνικών δικτύων και στις γνωσιακές επιστήμες.

6.2 Μελλοντικές κατευθύνσεις

Η συζήτηση μεταξύ του Roger Penrose και του Paul Churchland, καθώς και η εξερεύνηση της φύσης της συνείδησης, εξακολουθεί να είναι ένα ζωντανό και εξελισσόμενο πεδίο έρευνας. Κοιτάζοντας μπροστά, εμφανίζονται διάφορες μελλοντικές κατευθύνσεις που υπόσχονται να προωθήσουν την κατανόησή μας για τη συνείδηση και τους υποκείμενους μηχανισμούς της. Μια μελλοντική κατεύθυνση είναι η ενσωμάτωση πολλαπλών θεωρητικών προοπτικών. Αναγνωρίζοντας την πολύπλοκη φύση της συνείδησης, οι ερευνητές μπορούν να επωφεληθούν από το συνδυασμό ιδεών από διαφορετικές προσεγγίσεις, όπως υπολογιστικές, νευροεπιστημονικές, κβαντικές και φιλοσοφικές. Με την ενσωμάτωση αυτών των διαφορετικών απόψεων, οι ερευνητές μπορούν να αναπτύξουν πιο ολοκληρωμένες θεωρίες που αντιπροσωπεύουν διάφορες πτυχές των συνειδητών εμπειριών και γεφυρώνουν τα υπάρχοντα κενά στην κατανόηση του φαινομένου. Οι εξελίξεις στην τεχνολογία και τις εμπειρικές μεθόδους διανοίγουν μια πολλά υποσχόμενη οδό για τη μελλοντική έρευνα. Οι τεχνικές απεικόνισης του εγκεφάλου, όπως η λειτουργική απεικόνιση μαγνητικού συντονισμού (fMRI) και η ηλεκτροεγκεφαλογραφία (EEG), συνεχίζουν να βελτιώνονται όσον αφορά τη χωρική και χρονική ανάλυση. Αυτές οι τεχνολογίες επιτρέπουν στους ερευνητές να μελετήσουν τις νευρικές συσχετίσεις της συνείδησης με μεγαλύτερη ακρίβεια και να διερευνήσουν τη δυναμική της εγκεφαλικής δραστηριότητας κατά τη διάρκεια διαφορετικών καταστάσεων συνείδησης. Επιπλέον, καινοτόμες μέθοδοι όπως η οπτογενετική και η βαθιά εγκεφαλική διέγερση μας δίνουν τη δυνατότητα για παρεμβάσεις στη νευρική δραστηριότητα και για διερεύνηση των αιτιωδών σχέσεων μεταξύ των εγκεφαλικών διεργασιών και των συνειδητών εμπειριών.

Οι διεπιστημονικές συνεργασίες θα διαδραματίσουν κρίσιμο ρόλο στην προώθηση της κατανόησης της συνείδησης. Οι συνεργασίες επιστημόνων, φιλοσόφων, ψυχολόγων, νευρολόγων, επιστημόνων υπολογιστών, και άλλων σχετικών κλάδων μπορούν να προωθήσουν γόνιμες ανταλλαγές ιδεών, μεθοδολογιών και δεδομένων. Συνδυάζοντας την εμπειρία και τις προοπτικές από διαφορετικούς τομείς, οι ερευνητές μπορούν να αντιμετωπίσουν την πολυπλοκότητα του φαινομένου από πολλαπλές οπτικές γωνίες

οικοδομώντας νέες έννοιες και εννοιολογικά περιεχόμενα, κάτι ενδεχομένως αδύνατο μέσα σε μια ενιαία πρακτική.

Η ανάπτυξη πιο εξελιγμένων υπολογιστικών μοντέλων θα δώσει ισχυρή ώθηση στη μελλοντική έρευνα για τη συνείδηση. Ο τομέας της τεχνητής νοημοσύνης και της μηχανικής μάθησης συνεχίζει να αναπτύσσεται με ταχείς ρυθμούς, επιτρέποντας τη δημιουργία όλο και πιο πολύπλοκων, ρεαλιστικών υπολογιστικών μοντελοποιήσεων γνωσιακής λειτουργίας. Αυτά τα μοντέλα μπορούν να προσομοιάσουν διάφορες πτυχές των συνειδητών εμπειριών και να συμβάλουν στην κατανόηση της σχέσης μεταξύ των διαδικασιών του εγκεφάλου και της συνείδησης. Η ενσωμάτωση υπολογιστικών μοντέλων στα εμπειρικά δεδομένα και σε ένα αυστηρά θεμελιωμένο θεωρητικό πλαίσιο μπορεί να διευκολύνει την εξερεύνηση διαφορετικών υποθέσεων και να βοηθήσει στην ανάπτυξη πιο εκλεπτυσμένων θεωριών συνείδησης.

Οι εξελίξεις στις κβαντικές τεχνολογίες και η κατανόησή μας για τα κβαντικά φαινόμενα μπορεί να παρέχουν νέες γνώσεις σχετικά με τη φύση της συνείδησης. Κβαντικές προσεγγίσεις συνεχίζουν να εξετάζονται και να βελτιώνονται, ενώ η μελλοντική έρευνα μπορεί να ρίξει φως στις επιπτώσεις των κβαντικών επιδράσεων στις συνειδητές εμπειρίες.

Η ιστορική φάση στην οποία βρίσκεται η επιστημονική, αλλά και συνολικά, η ανθρώπινη κοινότητα επιβάλλει τη συστηματική ενασχόληση με τις προβληματικές που τέθηκαν συνολικά στην παρούσα εργασία. Υπό το φως των νέων τεχνολογικών εξελίξεων, καλούμαστε να εξοικειωθούμε με ολοένα και περισσότερες - συχνά θολές και δυσνόητες - θεωρητικές έννοιες αναπτύσσοντας τα απαραίτητα εργαλεία για την προσαρμογή μας στον σύγχρονο, διαρκώς μεταβαλλόμενο περιβάλλον. Η μελλοντική πορεία των μελετών της συνείδησης οφείλει να περιλαμβάνει - ανάμεσα σε άλλα - την ενσωμάτωση πολλαπλών θεωρητικών προοπτικών, την αξιοποίηση των εξελίξεων στην τεχνολογία και τις εμπειρικές μεθόδους, την προώθηση διεπιστημονικών συνεργασιών, την ανάπτυξη εξελιγμένων υπολογιστικών μοντέλων, την εξερεύνηση κβαντικών φαινομένων και τη διερεύνηση ασυνήθιστων συνειδησιακών καταστάσεων. Αυτές οι κατευθύνσεις έχουν τη δυνατότητα να εμβαθύνουν την κατανόησή μας για τη συνείδηση και να ωθήσουν την επιστημονική σκέψη προς τα εμπρός, προσδοκώντας να διαφωτίσουν μια από τις πιο ενδιαφέρουσες και αινιγματικές πτυχές της ανθρώπινης ύπαρξης.

Κατάλογος Εικόνων

ΕΙΚΟΝΑ 1: ΜΙΑ ΑΠΕΙΚΟΝΙΣΗ ΤΗΣ ΔΥΝΑΤΟΤΗΤΑΣ ΠΟΛΛΑΠΛΗΣ ΠΡΑΓΜΑΤΩΣΗΣ.	20
ΕΙΚΟΝΑ 2: ΠΑΡΤΙΔΑ ΣΚΑΚΙ.....	34
ΕΙΚΟΝΑ 3: ΈΝΑΣ ΤΥΠΙΚΟΣ ΝΕΥΡΩΝΑΣ.....	43
ΕΙΚΟΝΑ 4: ΟΙ ΔΙΑΔΟΧΙΚΟΙ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΙ ΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΠΡΟΤΥΠΩΝ (PATTERNS) ΕΝΕΡΓΟΠΟΙΗΣΗΣ.....	44
ΕΙΚΟΝΑ 5: ΔΙΚΤΥΟ ΕΠΑΝΑΤΡΟΦΟΔΟΤΗΣΗΣ	46
ΕΙΚΟΝΑ 6: Η ΦΛΟΙΩΔΗΣ ΔΡΑΣΤΗΡΙΟΤΗΤΑ	49
ΕΙΚΟΝΑ 7: ΟΙ ΑΞΟΝΙΚΕΣ ΠΡΟΒΟΛΕΣ ΠΟΥ ΣΥΝΔΕΟΥΝ ΟΛΕΣ ΤΙΣ ΠΕΡΙΟΧΕΣ ΤΟΥ ΕΓΚΕΦΑΛΙΚΟΥ ΦΛΟΙΟΥ ΜΕ ΤΟΝ ΕΝΔΟΠΕΤΑΛΙΟ ΠΥΡΗΝΑ ΤΟΥ ΘΑΛΑΜΟΥ... ..	50
ΕΙΚΟΝΑ 8: (Α) ΚΑΘΕΤΗ ΠΕΡΙΠΤΩΣΗ. (Β) ΚΛΙΣΗ 45 ΜΟΙΡΩΝ. (C) ΚΛΙΣΗ 60 ΜΟΙΡΩΝ	56
ΕΙΚΟΝΑ 9: ΕΛΕΥΘΕΡΗ ΠΤΩΣΗ: (Α) ΚΑΘΕΤΗ ΠΕΡΙΠΤΩΣΗ. (Β) ΜΙΚΡΗ ΣΥΝΙΣΤΩΣΑ ΟΜΟΙΟΜΟΡΦΗΣ ΟΡΙΖΟΝΤΙΑΣ ΚΙΝΗΣΗΣ. (C) ΜΕΓΑΛΥΤΕΡΗ ΣΥΝΙΣΤΩΣΑ ΟΜΟΙΟΜΟΡΦΗ ΟΡΙΖΟΝΤΙΑΣ ΚΙΝΗΣΗΣ	57
ΕΙΚΟΝΑ 10: ΣΩΜΑΤΑ ΠΟΥ ΠΕΦΤΟΥΝ ΕΛΕΥΘΕΡΑ ΣΕ ΑΝΟΜΟΙΟΜΟΡΦΟ ΒΑΡΥΤΙΚΟ ΠΕΔΙΟ.....	57

Βιβλιογραφία

- Miller, V., & Miller, R. (1982). *René Descartes: Principles of Philosophy*.
- Churchland, P. M. (1996). *The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain*. The MIT Press.
- Crane, T. (2001). *History of the Mind Body Problem*. Routledge.
- Curley, E. (1994). *A Spinoza Reader: The Ethics and Other Works*. Princeton University Press.
- Deleuze, G. (1981). *Spinoza. Philosophie pratique*. Paris: Les Editions de Minuit.
- Dennett, D. (2004). Can Machines Think? Στο *Alan Turing: Life and Legacy of a Great Thinker*. Springer.
- Descartes, R. (2008). *Meditations on First Philosophy*. Oxford University Press.
- Koch, C., & Tononi, G. (2008, June). Can Machines Be Conscious? *Institute of Electrical and Electronics Engineers, Spectrum*.
- Nagel, T. (1974, October). What Is It Like to Be a Bat? *The Philosophical Review*, pp. 435-450.
- Penrose, R., & Hameroff, S. (1996). Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness. *Mathematics and Computers in Simulation*.
- Penrose, R. (1996). *Shadows of the Mind: A Search for the Missing Science of Consciousness*. Oxford University Press.
- Place, U. (1956). Is consciousness a brain process? *The British Journal of Psychology*.
- Putnam, H. (2010). Philosophy and our mental life. Στο H. Putnam, *Mind, Language and Reality*. Cambridge University Press.
- Putnam, H. (2010). The Nature of Mental States. Στο H. Putnam, *Mind, Language and Reality*. Cambridge University Press.
- Ryle, G. (1949). *The Concept of Mind*. University of Chicago Press.
- Searle, J. (1980, September). Minds, brains, and programs. *Behavioral and Brain Sciences*, pp. 417 - 424.
- Searle, J. (2002). *Consciousness and Language*. Cambridge University Press.
- Stanford University. (1997, 9 22). Retrieved from <https://plato.stanford.edu/entries/leibniz-mind/>.
- Μαραγκός, Ν. (1995). Η συνείδηση, φυσικά. *Δευκαλίων*.

Μπαλάς, Α. (2022). *Ξεφλουδίζοντας πατάτες ή λειαίνοντας φακούς*. Πανεπιστημιακές Εκδόσεις Κρήτης.