

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

## Standardized rules for collation, representation, and coding of data in the CoCoMac database

([www.cocomac.org](http://www.cocomac.org))

### Table of contents

<b>A. GENERAL INFORMATION .....</b>	<b>5</b>
A.1 ADMINISTRATION GUIDELINES .....	5
A.2 GENERAL ORDER OF ENTERING DATA.....	6
<b>B. BIBLIOGRAPHIC DATA .....</b>	<b>7</b>
<i>B.1 Literature .....</i>	<i>7</i>
<i>B.2 Literature_Abbreviations_Journals.....</i>	<i>10</i>
<i>B.3 Journals / Chapters / Books.....</i>	<i>10</i>
<i>B.4 Authors.....</i>	<i>11</i>
<i>B.5 Literature_LinkTable .....</i>	<i>11</i>
<i>B.6 Administration_Proofreading tables .....</i>	<i>11</i>
<b>C. MAPPING DATA .....</b>	<b>13</b>
C.1 GENERAL PRINCIPLES OF PARCELLATION-BASED DATA REPRESENTATION IN CoCoMac.....	13
<i>C.1.1 The concepts of BrainSites and BrainMaps .....</i>	<i>13</i>
<i>C.1.2 Delineation and adoption of BrainSites .....</i>	<i>16</i>
<i>C.1.3 When are BrainMaps and BrainSites defined in CoCoMac? .....</i>	<i>18</i>
C.2 TWO SPECIAL BRAINMAPS.....	20
C.2.1 <i>Acronym Map (AM) .....</i>	<i>21</i>
C.2.2 <i>General Map (GM) .....</i>	<i>25</i>
C.2.3 <i>Comparing AM and GM .....</i>	<i>27</i>

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

C.3	DEFINITION OF DIFFERENT TYPES OF BRAINSITES.....	28
C.4	TABLES THAT CONTAIN MAPPING DATA .....	30
C.4.1	<i>BrainMaps_BrainSites_SiteClasses</i> .....	30
C.4.2	<i>BrainMaps_BrainSites_SiteTypes</i> .....	31
C.4.3	<i>BrainMaps</i> .....	33
C.4.4	<i>BrainMaps_Delineation</i> .....	34
C.4.5	<i>BrainMaps_UnspecificAdoption</i> .....	34
C.4.6	<i>BrainMaps_Methods</i> .....	35
C.4.7	<i>BrainMaps_BrainSiteAcronyms</i> .....	35
C.4.8	<i>BrainMaps_BrainSites</i> .....	36
C.4.9	<i>InterMapRelations</i> .....	38
C.4.10	<i>InterMapRelations_References</i> .....	40
<b>D.</b>	<b>EXPERIMENTAL DATA.....</b>	<b>42</b>
D.1	TABLES CONTAINING EXPERIMENTAL DATA.....	42
D.1.1	<i>Methods</i> .....	42
D.1.2	<i>Methods_Animals</i> .....	43
D.1.3	<i>Methods_Animals_Details</i> .....	44
D.1.4	<i>Methods_RadioactiveTracers</i> .....	44
D.1.5	<i>Injections</i> .....	45
D.1.6	<i>Injections_AffectedSites</i> .....	46
D.1.7	<i>Injections_Laminae</i> .....	47
D.1.8	<i>LabelledSites_Descriptions</i> .....	48
D.1.9	<i>LabelledSites_Data</i> .....	48
D.1.10	<i>LabelledSites_Data_Laminae</i> .....	49
D.1.11	<i>LabelledSites_Data_Sublaminae</i> .....	50
D.1.12	<i>LabelledSites_Data_QD_Site_Nbr</i> .....	51
D.1.13	<i>LabelledSites_Data_QD_Site_Percent</i> .....	51
D.1.14	<i>LabelledSites_Data_QD_Lamina_Percent</i> .....	51
<b>E.</b>	<b>OTHER TABLES .....</b>	<b>53</b>
E.1	TABLES ONLY USED BY ORT ALGORITHMS .....	53
E.2	BASE TABLES (I.E. WITHOUT FOREIGN KEYS) .....	53

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

E.3	ADMINISTRATION TABLES.....	54
<b>F.</b>	<b>GENERAL CODING RULES.....</b>	<b>55</b>
F.1	CHOOSING APPROPRIATE TYPES OF BRAINSITES.....	55
F.2	IDENTIFICATION OF (UN)LABELLED BRAIN SITES.....	56
F.3	EXTENSION CODES (ECs).....	57
F.4	DENSITY CODING.....	59
F.5	LAMINAR PATTERNS.....	60
F.6	SPECIAL CONSIDERATIONS.....	60
<b>G.</b>	<b>PRECISION OF DESCRIPTION CODES (PDCS).....</b>	<b>62</b>
G.1	IDENTIFICATION OF BRAIN SITES.....	65
G.2	EXTENSION CODES.....	68
G.3	DENSITY OF LABEL.....	72
G.4	LAMINAR PATTERNS.....	74
G.5	RELATIONS BETWEEN BRAINMAPS.....	76
<b>H.</b>	<b>GUIDELINES FOR PROOFREADING PREVIOUSLY ENTERED DATA 82</b>	
H.1	GENERAL REMARKS.....	82
H.2.	BIBLIOGRAPHIC DATA.....	82
H.2.1	<i>Literature</i> .....	82
H.2.2	<i>Literature_LinkTable</i> .....	82
H.3	MAPPING DATA.....	83
H.3.1	<i>BrainMaps</i> .....	83
H.3.2	<i>BrainMap_BrainSiteAcronyms</i> .....	83
H.3.3	<i>BrainMaps_BrainSites</i> .....	83
H.3.4	<i>BrainMaps_Methods</i> .....	83
H.3.5	<i>InterMapRelations</i> .....	83
H.3.6	<i>InterMapRelations_References</i> .....	83
H.4	EXPERIMENTAL DATA.....	84
H.4.1	<i>Injections</i> .....	84
H.4.2	<i>Injections_Laminae</i> .....	84

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

<i>H.4.3</i>	<i>LabelledSites_Data</i> .....	84
<i>H.4.4</i>	<i>LabelledSites_Data_Laminae</i> .....	85
<i>H.4.5</i>	<i>LabelledSites_Descriptions</i> .....	85
<b>I.</b>	<b>LIST OF ABBREVIATIONS AND TECHNICAL TERMS</b> .....	<b>86</b>

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

## A. General information

### A.1 Administration guidelines

- For all technical aspects of the work with CoCoMac (e.g. setup, locations of master and client databases, backup regulations) see the technical manual for CoCoMac (manual\_technicalSetup\_cocomac).
- Before working with CoCoMac, make sure you have disabled **all** options in the Tools/AutoCorrect menu of Access, otherwise you may be unpleasantly surprised about unexpected changes to the names of brain structures...
- Document every main step of your work in CoCoMac's log file (table "Administration\_Logtable").
- Concerning experimental data, we restrict the scope of CoCoMac with regard to methods, species and age of animals: only data from *tracing studies on adolescent or adult Macaque monkeys* is accepted. As for mapping data (definitions of parcellation schemes, relations), all articles are accepted independent of their methodology as long as they refer to *adolescent or adult Macaque monkeys*
- Read the article entirely before entering any kind of data. Do not omit introductions or discussion - they may contain important information about relations between brain maps.
- The relational-hierarchical structure of CoCoMac requires that data are entered in a certain sequence: First, enter bibliographic data, then mapping data, finally experimental data - the text below indicates the exact order of the tables.
- For each table, make sure that you have entered all data in hierarchically higher tables that the foreign keys of your current table refer to (otherwise Access will tell you...).
- In case of contradictions between information delivered by text/tables and figures, the textual information is treated as superior. Report contradictions

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

in the field “Comments” and choose the appropriate PD-Code. **NOTE:** It is not necessarily a contradiction if a brain site is stated to be labelled by the text, but sections shown by the figures display no label for this particular brain site. This does not preclude the possibility that other sections, which are not shown by the figures, would confirm the label for this particular brain site.

## A.2 General order of entering data

1. Read the entire article.
2. Enter the **bibliographic data**.
3. **Mapping data:** first, determine how the article defines brain structures to represent its data: Does the article use delineation and / or unspecific adoption? If so, enter data for  
    “BrainMaps\_BrainSiteAcronyms” (if new abbreviations are introduced),  
    “BrainMaps”,  
    “BrainMaps\_Delineation” (if necessary),  
    “BrainMaps\_UnspecificAdoption” (if necessary),  
    “BrainMaps\_Methods” (if structures have been delineated),  
    “BrainMaps\_BrainSites”  
(in this order).  
Defining the brain map and brain sites of the article is very important - the representation of all experimental data depends on their proper definition!
4. Enter the information of relations between different parcellation schemes as given by text and/or figures into “InterMapRelations” and “InterMapRelations\_References”.
5. **Experimental data:** Enter data for “Methods” (and associated tables), “Injections”, “LabelledSites\_Descriptions”, “LabelledSites\_Data” and associated tables (quantitative data, laminar patterns).

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

## **B. Bibliographic data**

### **B.1 Literature**

This table is the highest in the hierarchy of the whole database. It contains basic information about the articles contained by CoCoMac and provides a unique ID for them.

- **ID:** This is composed of the first letters of the authors' surnames plus two digits for the year of print. If there are more than 6 authors, concatenate only the first 6 letters; add "al" (for "alii") and then the two digits.

**NOTE:** If you enter two papers with the same ID, add a lowercase letter as suffix to their IDs (e.g. CP95a, CP95b). Should a new paper be entered whose ID matches that of another paper which has already been entered previously, do not change the ID of the entered paper, i.e. do not add an "a" to it. This would require to change dependent entries in other tables (e.g. IDs of injections and IDs of brain sites in "BrainMaps\_BrainSites") which is quite an error-prone process. Instead, just add the suffix "b" to the new paper and leave the ID of the previously entered paper as it is.

Primary key.

- **Title:** Full title of the article.
- **Year:** Year of print.
- **Journal / Chapter / Book:** This field contains information whether the paper is an article from a Journal (J), an entire Book (B), or a Chapter from a book (C).
- **Abstract:** If available in MEDLINE, copy the abstract and paste it into this field. OPTIONAL.
- **Status\_DataEntry\_TracingData:** This field indicates the degree of completeness by which experiment-related data (tracing methods, injections, labeled sites) described by this article have been entered so far. Options include: (1) "?": status of data entry is unclear. (2) "no data provided": the article does not report any tracing data. (3) "not started":

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

the article does contain tracing data, but it has not been started to enter them. (4) "started (but not completed)" data entry has been started, but not completed. (5) "completed": all tracing data described by the article have been entered.

This field is a foreign key to table "StatusOptions\_DataEntry".

- **Status\_DataEntry\_MappingData:** This field indicates the degree of completeness by which mapping data (i.e., BrainMaps, BrainSites, and InterMapRelations) described by this article have been entered so far. Options include: (1) "?": status of data entry is unclear. (2) "no data provided": the article does not report any mapping data. (3) "not started": the article does contain mapping, but it has not been started to enter them. (4) "started (but not completed)": data entry has been started, but not completed. (5) "completed": all mapping data described by the article have been entered.

This field is a foreign key to table "StatusOptions\_DataEntry".

- **Status\_Proofreading\_TracingData:** This field states to what degree the CoCoMac representation of experiment-related data (tracing methods, injections, labeled sites) described by this article has been proofread so far. Options include: (1) "?": proofreading status is unclear. (2) "not applicable": the article does not report any tracing data therefore proofreading can't be applied. (3) "not started": the article does contain tracing data, but proofreading has not started yet. (4) "started (but not completed)": data entry has been started, but not completed. (5) "completed": all tracing data described by the article have been entered.

This field is a foreign key to table "StatusOptions\_Proofreading".

- **Status\_Proofreading\_MappingData:** This field states to what degree the CoCoMac representation of mapping data (i.e., BrainMaps, BrainSites, and InterMapRelations) described by this article has been proofread so far. Options include: (1) "?": proofreading status is unclear. (2) "not applicable": the article does not report any mapping data therefore proofreading can't be applied. (3) "not started": the article does contain mapping data, but proofreading has not started yet. (4) "started (but not

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

completed)": data entry has been started, but not completed. (5)

"completed": all mapping data described by the article have been entered.

This field is a foreign key to table "StatusOptions\_Proofreading".

**NOTE:** For the two StatusOptions fields, there are 7 variants of option 5 (i.e. completed proofreading) distinguishing whether proofreading has been performed by (i) the same person who entered the data, (ii) by an independent CoCoMac data collator, (iii) by one of the original authors. or (iv) by several of them. For a given article, there may well exist several independent processes of proofreading (that should be temporally non-overlapping, however). For example, the data representation of an article in CoCoMac might be proofread first by the same person who entered the data, then by an independent CoCoMac data collator and finally by one of the original authors. While each of these processes should be documented in the tables "Administration\_Proofreading\_TracingData" and "Administration\_Proofreading\_MappingData", respectively, the "Status\_Proofreading\_TracingData" "Status\_Proofreading\_MappingData" fields in the "Literature" table should always refer to the hierarchically highest (i.e. giving the highest degree of independent control) AND completed\_process. For example, assume that the tracing data of a given article have already been completely proofread by the same person who entered them and are currently being proofread by one of the original authors. This would be represented in CoCoMac in the following way: the "Administration\_Proofreading\_TracingData" table would contain one entry "completed (by same data collator)" that is linked to the name of the data collator and another entry "started (but not completed)" that is linked to the name of the original author. The "Status\_ProofreadingData" field in the table "Literature" would then state "completed (by same data collator)" the because this is the only completed proofreading process. When the original author has finished proofreading the data, this statement in the "Literature" table becomes "completed (by same data collator) and double-checked (by original author)".

**NOTE:** Whenever you finished a session of data entry, make sure to

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

update the information contained by the "Status\_DataEntry" fields in the "Literature" table.

**NOTE:** Whenever you start or complete a proofreading process, make sure to update (i) the information contained by the "Status\_Proofreading\_TracingData" and "Status\_Proofreading\_MappingData" fields in the "Literature" table and (ii) the information contained by the "ProofreadingStatus" field in the "Administration\_Proofreading\_TracingData" and "Administration\_Proofreading\_MappingData" tables, respectively.

- **Physical copy:** Tick this box if a physical copy of the article has been stored in the archive.
- **Comments:** OPTIONAL.
- **dbCollator:** Initials of the database collator who created this entry or performed last changes to it.

## B.2 Literature\_Abbreviations\_Journals

This table consists of only one field that stores the abbreviations that are commonly used to refer to journals, e.g. "J. Comp. Neurol." for the "Journal of Comparative Neurology".

## B.3 Journals / Chapters / Books

This table contains specific information on the article, depending on whether it is an article, a chapter, or a book.

- **Editoring Authors** in "Chapters": List the editors of the book in the form <surname\_author1> <initials\_firstname\_author1>, <surname\_author2> <initials\_firstname\_author2>,...
- **Journal** in "Journals": Enter the abbreviation that is stored by "Abbreviations\_Journals", e.g. "J. Comp. Neurol." for the "Journal of Comparative Neurology". Foreign key, linked to "Abbreviations\_Journals".

All other fields of these 3 tables should be self-explaining.

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

## B.4 Authors

This table lists all authors of the papers known by CoCoMac.

- **LastName\_Author:** Last name of the author. Part of primary key.
- **Initials\_FirstName\_Author:** Initials of the first name of the author. Part of primary key.

## B.5 Literature\_LinkTable

This table links each paper with all authors who have contributed to it.

- **ID\_Literature:** ID of the paper. Part of primary key. Foreign key to "Literature".
- **LastName\_Author:** Last name of the author. Part of primary key. Foreign key to "Authors".
- **Initials\_FirstName\_Author:** Initials of the first name of the author. Part of primary key. Foreign key to "Authors".
- **Position:** position of the author with respect to the author list of the paper (i.e. is the author the first, second, etc. author).

A convenient way to enter data for this table is to use the form "AddLink\_Literature".

## B.6 Administration\_Proofreading tables

This paragraph describes two tables with identical fields:

"Administration\_Proofreading\_TracingData" and

"Administration\_Proofreading\_MappingData". These tables document which of the data that had been entered into CoCoMac has already been proofread, at what date, to what degree, and by whom.

- **ID\_Literature:** ID of the proofread paper. Part of primary key. Foreign key to "Literature".
- **Date:** Date of proofreading (YY-MM-DD). Part of primary key.
- **ProofreadingStatus:** how complete is the proofreading process? Part of primary key. Foreign key to table "StatusOptions\_Proofreading".

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

- **Comment:** OPTIONAL
- **LastName:** last name of the data collator or the original author who performed the proofreading. Part of primary key. Foreign key to table "Literature\_Authors".
- **Initials:** initials of the data collator or the original author who performed the proofreading. Part of primary key. Foreign key to table "Literature\_Authors".

**NOTE:** For each article, this table can contain more than one entry when several people have proofread the article. In contrast, the "Status\_Proofreading\_TracingData" "Status\_Proofreading\_MappingData" fields in the "Literature" table always give a single statement on the proofreading status, i.e. they refer to the hierarchically highest (i.e. giving the highest degree of independent control) AND completed proofreading process for a given article. See section B.1 for further details.

## **C. *Mapping data***

### **C.1 General principles of parcellation-based data representation in CoCoMac**

#### **C.1.1 The concepts of BrainSites and BrainMaps**

As already mentioned above, tracing studies have so far not described their results by coordinates in Euclidean 3D space, but instead refer to descriptive entities which are defined by microstructural (e.g. cytoarchitectonics, receptor or enzyme distributions) or functional (e.g. electrophysiological) properties. These units are what we will generally refer to as "brain sites" (more specifically, they are usually called "brain sites", "areas" or "fields" in the cortex, and "nuclei" in the subcortex). A set of such microstructurally-functionally defined entities as presented by a particular article constitutes a "parcellation scheme" or "brain map". In analogy to a spatial description of brain data in Euclidean 3D-space, representing brain data on the basis of parcellation schemes can be envisaged as a non-Euclidean representation in a multidimensional space of microstructural-functional properties. For example, in a cytoarchitectonic study which defines areal boundaries on the basis of (i) relative width of supra- and infragranular layers, (ii) staining intensity of layer IV, (iii) density of pyramidal cells in layer III, (iv) density of pyramidal cells in layer V, these criteria define the four dimensions of the non-Euclidean feature space in which the delineated brain sites would be located. Unfortunately, due to a variety of applied methods and the high observer-dependency of many microstructural-functional criteria, there exists a large number of parcellation schemes that differ in nomenclature, number, location, size, boundaries, and representation of the delineated brain sites. This variability is especially high for the cortex, but can also be found for subcortical structures (e.g. thalamus, amygdala, brain stem). The combination

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

of different methods and observer-dependent criteria can make comparing data described by two different parcellation schemes substantially difficult. Formally, such a comparison is equivalent to comparing two points from two non-Euclidean spaces whose dimensions are characterized by the microstructural-functional criteria of the two parcellation schemes. Depending on potential differences between these criteria, the two feature spaces thus have different numbers and/or different types of dimensions. The question how difficult it is to compare data based on two parcellation schemes can thus be formally rephrased as "Is there a determinable mapping between the two spaces?" or, less formally, "Is there a way to relate (and maybe match) the two parcellations schemes so that they may be compared in a meaningful way?".

Let's look at an example of several parcellations of premotor cortex. For the two studies of Brodmann (1909) and von Bonin & Bailey (1947), the experimental method (i.e. Nissl staining) and the criteria used for delineation (i.e. certain cytoarchitectonic properties determined by visual inspection) are very similar, but of relatively high observer-dependency. In spite of the resulting obvious differences between these two maps, one can still compare these parcellations relatively easily because the dimensions of the two microstructural spaces are of similar quality. One would find, for example, that the general cytoarchitectonic features described by Brodmann for his area 6 (e.g. agranular cortex without Betz cells in layer V) are still valid for each of the premotor areas FB, FBA, and FCBm recognized by von Bonin & Bailey, but that the latter used additional criteria to distinguish these subdivisions within Brodmann's area 6. A third study by Vogt & Vogt (1919) also follows a classical architectonic approach, using visually determined changes in myeloarchitecture to determine areal borders. If this parcellation is compared to either Brodmann (1909) or von Bonin & Bailey (1947), how can the microstructural spaces be related to each other as their dimensions are of an entirely different quality? In other words, how can we determine whether a border determined by purely myeloarchitectonic criteria matches an areal border that has been determined by purely cytoarchitectonic criteria? This

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

direct comparison is only possible if there is a way of establishing a clear match between particular myeloarchitectonic and particular cytoarchitectonic features - studies combining both criteria (e.g. Barbas & Pandya 1987) can indeed serve as a connecting link that allows such a comparison. Beyond mere architectonic criteria, the dimensional mismatch between the feature spaces of parcellations becomes even larger if any of the structural studies above is compared to other studies of premotor cortex that *only* apply electrophysiological criteria to distinguish brain sites on the basis of different response properties (Luppino et al. 1991; Mitz & Wise 1986).

It seems that in such cases where the microstructural-functional feature spaces of two parcellations are not directly comparable, additional information is needed which must arise from the parcellations themselves, not from their underlying criteria. Commonly, this is topographic information that results from a comparison of the two parcellations with respect to relative size and position of brain sites and relative location to landmarks. This process certainly is highly observer-dependent and additionally problematic due to the high inter-individual differences between brains. Nevertheless, rough topographic information is always obtainable for two given parcellation schemes and may be the only way to perform any comparison at all for parcellations with no direct comparability of their microstructural-functional feature spaces.

Both approaches - comparisons by microstructural-functional and topographic criteria - are frequently found in the literature. Unfortunately, very few authors explicitly point out the basis of their statements on relations between different brain maps. Often, a combination of both microstructural-functional and topographic comparisons seems to underlie such statements.

### **What are the conclusions for CoCoMac?**

#### 1. *Representation of brain sites:*

Delineated brain sites are the most central objects in CoCoMac.

Evaluating any given article, a crucial first step is to carefully represent its defined or adopted brain sites. At a later stage, this allows to represent experimental data in their original nomenclature, thus minimizing

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

interpretation.

Due to their very coarse resolution and variations of inter-individual brain morphology, topographic entities (e.g. sulci, gyri, lobes) without references to microstructural or functional criteria are not represented.

## 2. *Representation of comparisons/relations between brain maps:*

Often, we often neither know the basis (microstructural-functional vs. topographic criteria, assumptions) nor the quality of the comparisons between parcellations - therefore all comparative statements are equally represented, but need to be carefully cited for documentation.

## 3. *Mapping data between different brain maps:*

During the process of data collation and representation, data are not mapped but described on the basis of their original parcellation scheme (see above). Mapping experimental data between different parcellations is done algorithmically by ORT, based on the comparisons between different parcellations as stated in the literature. The mapping process is thus reproducible and transparent as all underlying procedures (formally defined principles of data processing) and assumptions (statements about relations between parcellations) are clearly documented.

### **C.1.2 Delineation and adoption of BrainSites**

Finally, it must be explained that there are four different general ways how a particular article can define or adopt brain sites, which it needs to describe its results. Not all of these ways lead to the creation of brain sites that need to be represented within CoCoMac. The four possibilities are the following:

(A) ***Delineation***: The article experimentally defines locations and boundaries of the brain sites it refers to. The parcellation may either be a new one or confirm existing ones. In any case, the parcellation is clearly defined and described as to be a potential reference for future articles: the definitions of the brain sites are either given by textual descriptions of the delineated brain sites (e.g. extent, location, architectonic properties of the brain site) or by a figure (overview or sections) whose purpose clearly is to

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

demonstrate the locations and boundaries of brain sites (thus, not just any figure that shows experimental results projected upon delineated areas).

**NOTE:** It is sufficient if such a figure shows only one (and not all) borders of the brain site in question.

(B) ***Unspecific Adoption***: The article does not delineate brain sites itself, but adopts brain sites from one or several parcellation schemes. The relation between the brain sites and the parcellation schemes, however, is not unambiguously specified insofar as it is not stated what parcellation scheme a particular brain site is taken from:

- An areal name is used without any reference to a map at all (e.g. "...label was found in area 7..." or a figure showing an area 7 that is not explained any further).
- Several parcellations with similar / identical names for brain sites are referred to simultaneously without exact specification which definition applies to a particular brain site. Examples: "To describe our findings in the prefrontal cortex, we refer to standard parcellation schemes (Walker 1940; Carmichael & Price 1994; Preuss & Goldman-Rakic 1991)." or "We found cells in a region that has been identified as the SEF (Schlag and Schlag-Rey, 1987; Huerta and Kaas, 1990)." The parcellation schemes in these examples are difficult to distinguish by nomenclature as they contain similar and sometimes identical names, but may differ in their definition even for identically named areas (example: MLR91-24c is explicitly defined to be a subarea of VPR87-24c, in spite of the identical acronym – see Matelli et al. 1991, J. Comp. Neurol. Vol. 311, p. 460).

(C) ***Specific Adoption***: The article does not delineate brain sites itself, but adopts brain sites from one or various other parcellation schemes. In contrast to *Unspecific Adoption* (see above), the source for the adopted brain sites is clearly specified. Example 1: "We used the parcellation of Walker (1940) for the prefrontal areas and the map of Matelli et al. (1991)

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

for the premotor areas to describe our findings." Example 2: "To describe our results, we refer to the maps of Brodmann (1909) and von Bonin & Bailey (1947)." Although the statement in Example 2 seem not to be specific as such, the two maps addressed (i.e. B09 and BB47) use entirely different names to designate areas so that each reference to an area can be clearly assigned to one of the two maps on the basis of nomenclature alone.

**(D) Topographic description:** The article does not refer to a delineated brain site by its name or acronym, but (i) uses broad topographic labels to describe data OR (ii) displays data on figures which lack areal names and borders and for which none of the map(s) defined or adopted by the article is valid. Example: "Label was also found in the vicinity of the principal sulcus and in the dorsolateral prefrontal cortex."

These four general ways can also be combined, e.g. an article may delineate brain sites for one part of the cortex (option A) as well as adopt brain sites from another map for a different part of the cortex (options B or C). Also, an article which adopts brain sites only can clearly state the source map for some brain sites (option C) and be ambiguous for other brain sites (option B).

### C.1.3 When are BrainMaps and BrainSites defined in CoCoMac?

Within CoCoMac we explicitly define a BrainMap (i.e. create an entry in the table "BrainMaps") for any article that refers to a parcellation scheme with one of the following characteristics:

- The brain sites are clearly defined by the article: their location and boundaries are either described in detail by the text and/or the figures show a general map with clear boundaries (*Delineation*). Such brain sites - even if they only confirm existing parcellations - may be referred to by future articles and thus need to be explicitly represented within CoCoMac.

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

- The brain sites are adopted from other brain maps but their origin is not entirely clear (*Unspecific Adoption*). To explicitly document our interpretations we list these brain sites as members of the article's brain map and state the assumed relation to other maps in the table "InterMapRelations" (see PP99 as an example of an article where this case applies to many brain sites). Should future information reveal that our interpretations have been wrong, we can easily correct the respective relations without having to change the representation of the experimental data. On the other hand, if we directly used our interpretations for the representation of data in CoCoMac, we would violate the principle of transparency, and potential future corrections would require large and laborious changes.

Turned the other way round, we do *not* need to define a brain map in case of an article that has adopted all used brain sites from other maps in an unambiguous way so that we know exactly what map each brain site was taken from (*Specific Adoption*). Also, we do not include *topographic descriptions* into a brain map for an article. Instead, for describing the associated data, we choose a brain site defined by another map whose location approximately matches the topographical description. Example: for the vague description "...label was found in the vicinity of the principal sulcus..." we choose a suitably defined area, e.g. W40-46, and assign the PDC\_Site = P to express our assumption.

**NOTE:** Companion articles (i.e. if author X publishes two (or more) articles in the same issue of a particular journal) represent a special case. If any of these companion articles describes or presents a figure of a brain map, then our criteria for *Delineation* apply and we need to define these brain sites explicitly as members of the brain map of this article. If the article does not show or describe a map, it will often explicitly refer to another companion article, stating that its brain map is used (*Specific Adoption*). Sometimes, however, it simply uses the names of brain sites without giving clear references. Then you may assume that these brain sites are identical to the

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

ones used in the companion article(s). This situation is thus treated as a special case of *Specific Adoption*.

**NOTE:** Often articles define a particular brain site first in a general way and then subdivide it further (e.g. PP94 define an area 8A which is then subdivided into dorsal and ventral parts, i.e. areas 8Ad and 8Av). Subsequent articles may either refer to the superarea (8A) or to the subareas (8Ad and 8Av). Therefore, we must include all three definitions in the "BrainMaps\_BrainSites" table. Make sure, that you also add the appropriate relation in the "InterMapRelations" table (see below). In this example, you would have to state that PP94-8A contains the subareas PP94-8Ad and PP94-8Av (L-relation).

## C.2 Two special BrainMaps

The concept of BrainMaps and BrainSites as described in the paragraphs above is particularly suitable for brain regions where a variety of different histological and/or physiological definitions of areas and nuclei co-exist. Good examples of such regions in CoCoMac with parcellation-defined BrainSites are the cerebral cortex, the thalamus, and the amygdala.

In other parts of the brain, however, differences between parcellations are rare, either because of obvious demarcations of structures by macroscopic features (e.g. demarcation by white matter) or because concepts of fine-grained parcellations have not been considered important yet. In both cases, different authors refer to brain structures without explicit definitions, e.g. they refer to "the" caudate nucleus, "the" pons, "the" granular layer of "the" cerebellum etc. One may thus speak of generalised BrainSites.

During the development of CoCoMac, it has become obvious that each of these two different sets of BrainSites introduces a particular problem that requires some principles of data representation which are additional to the ones described in section C.1 above.

In the following, we will use the acronym "OM" (for "Original Maps") to refer either to the entire set of BrainMaps defined by original articles or to one

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

specific original BrainMap. In contrast, “AM” will denote the so-called “Acronym Map” and “GM” the so-called “General Map”.

## C.2.1 Acronym Map (AM)

One common problem is that in many brain maps BrainSites are used that have the same acronym and very probably the same meaning (i.e. delineation) but the literature provides no explicit statements on their identity. For example, articles describing connectivity of the visual cortex may mention an area “V1” without referring to any definition of this area (i.e. these are cases of unspecific adoption). Given the wide agreement on the definition of primary visual cortex, it is extremely likely that all these OM-BrainSites are identical. For each relation, one could state this assumption explicitly, using  $PDC\_Relation=P$  (see section G.5). This procedure has two disadvantages:

- (i) The user has to enter manually large numbers of relations that could be automatically generated by a simple rule: “Whenever there are two OM-BrainSites with identical acronyms and no explicit information on their relation, generate an I-relation for these BrainSites and define  $PDC\_Relation=P$  to express the underlying assumption.” Manual representation of these relations is not only tedious, but also error-prone.
- (ii) All these semantically equivalent OM-BrainSites are treated separately by the ORT algorithms although they could be effectively merged into one single BrainSite.

In this section, the concept of the AcronymMap (AM) is presented that tackles this problem. Defined briefly, AM is an algorithmically created set of BrainSites for each acronym in the table `BrainMaps_BrainSites_Acronyms` that is used by at least two BrainSites in CoCoMac. In other words, for each of these acronyms there is one element of AM that represents the acronym as a general BrainSite.

At first sight, this may appear to be a burden on computational efficiency as it adds a large number of additional BrainSites to the table `BrainMaps_BrainSites`. If used properly, however, AM considerably simplifies

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

the representation of BrainSites and relations and contributes to reducing the number of BrainSites that effectively enter the computations of ORT.

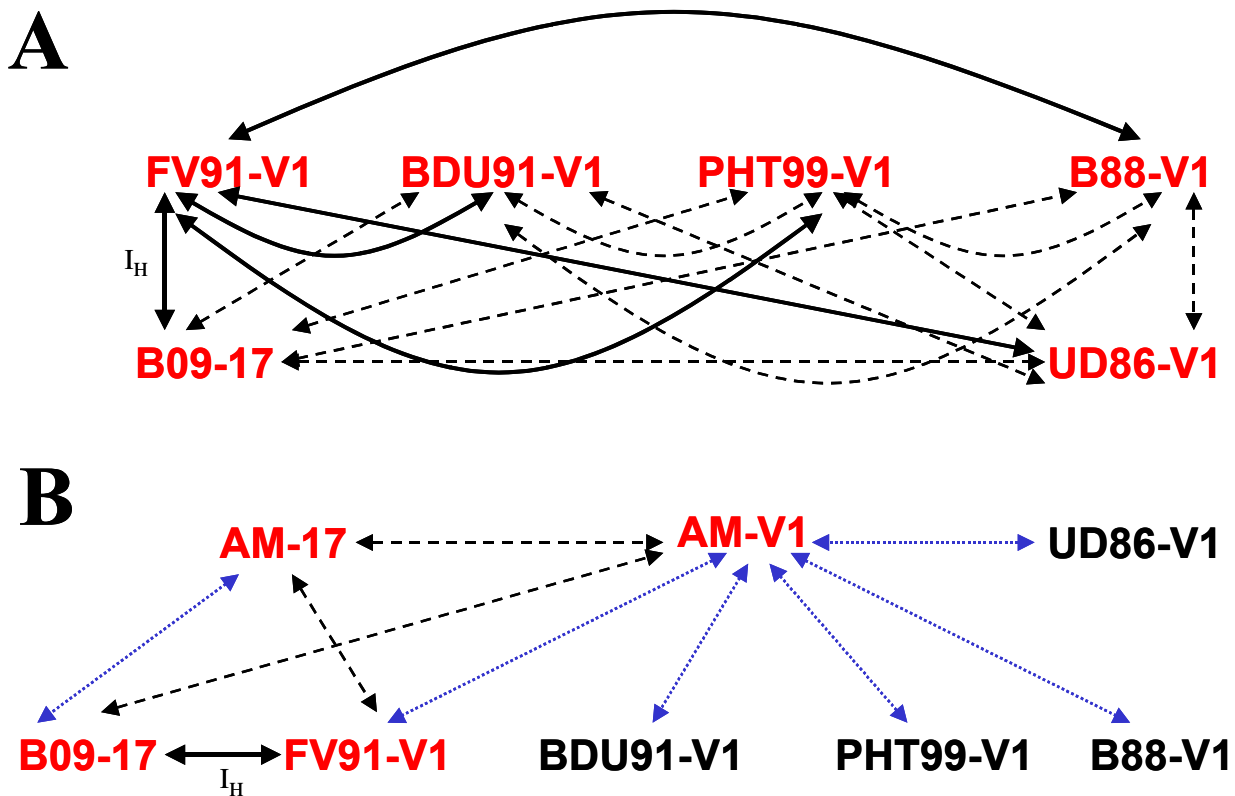
### **C.2.1.1 Definition and use of AM**

The user does not have to define the elements of AM nor to create relations to its elements. Instead, AM is generated anew algorithmically each time ORT is run. The following principles apply:

- For each acronym A in BrainMaps\_BrainSiteAcronyms that is used by at least two BrainSites, an element AM-A in BrainMaps\_BrainSites is created. If indices are present in acronyms, these are also included in the ID of AM-BrainSites. For example, for the four acronyms P#1 – P#4 (representing putamen, nucleus parafascicularis thalami, periamygdaloid cortex, and area prostriata, respectively), the corresponding elements of AM in BrainMaps\_BrainSites would be AM-P#1, ..., AM-P#4.
- For each element AM-A in the AcronymMap, I-relations between AM-A and all corresponding BrainSites  $M_i$ -A ( $M_i$  = BrainMap of article i, A=acronym) of BrainMaps\_BrainSites are created. These relations are assigned a new PDC (PDC\_Relation = R → this PDC denotes a relation automatically created on grounds of identical acronyms). For example, for the BrainSites BDU91-V1 and FV91-V1 the relations BDU91-V1 --I--> AM-V1 and FV91-V1 --I--> AM-V1 are created (see figure below). Note that no BrainSite is created in AM for an acronym that is used by only one BrainSite in BrainMaps\_BrainSites.
- Further relations between and to BrainSites of AM are automatically generated during the optimisation of the transformation path (see figure below). Placing the AM-specific PDC\_Relation (R) at the bottom of the PDC\_Hierarchy ensures that all relations entered by the data collators will override any relations that are created automatically via AM.

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany



This figure shows an example how 6 BrainSites of primary visual cortex are represented and processed without (A) and with (B) the AM procedure. Bold black arrows denote I-relations that need to be entered by the collator. Except for B09-17 --I-->FV91-V1 which is clearly stated in the literature (PDC\_Relation=H), these relations are assumed and thus PDC\_Relation=P. Dotted blue arrows represent I-relations with PDC\_Relation=R that are algorithmically inserted for BrainSites with identical acronyms but unknown relations. Dashed black arrows represent I-relations that are created automatically during the optimisation of the transformation path. Red BrainSites enter the optimisation of the transformation graph and all subsequent stages of ORT. Black BrainSites do not enter ORT. (A) Without AM, the data collator is required to enter at least 5 relations to ensure that transformation paths are eventually created between the 6 BrainSites all of which have to enter the graph optimisation routine. (B) Using AM, the data collator only needs to enter a single relation (the one that is documented in the literature), still all possible transformation paths are created. Another advantage compared to (A) is that only 4 BrainSites need to enter ORT.

AM thus summarises general notions how BrainSites are related to each other simply by nomenclature and not by underlying specific parcellations. One positive effect of this is that I-relations with PDC\_Relation = P between BrainSites with identical acronyms do not need to be represented by the user: all BrainSites with identical acronyms are automatically linked by I-relations to the same element of AM. During the optimisation of the transformation graph, these BrainSites will thus be linked directly by an I-relation as well – unless there is a different relation entered by the user that has a better PDC\_Relation (which will always be the case as long as PDC\_Relation=R is placed at the

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

bottom of the PDC\_Hierarchy). Also, using the automatic linkage via the AM prevents that mapping between identically named BrainSites fails.

### **C.2.1.2      *Implications for ORT***

By definition, all relations with PDC\_Relation = P express assumptions. If a given BrainSite OM-A only possesses such assumed relations to other BrainSites, we can omit OM-A from any ORT operations since the corresponding (and automatically created) BrainSite AM-A has the very same relations that are also based on assumptions. In contrast to initial expectations, however, the overall effect of AM on the number of BrainSites entering ORT computations does not seem to be positive: while a considerable number of OM-BrainSites can be replaced by AM-BrainSites, this reduction is outweighed by the fact that all elements of AM are added to the transformation path to ensure mapping between identically named BrainSites.

### **C.2.1.3      *Which problem does AM not solve?***

Since AM equates identity of BrainSites with identical acronyms, it cannot recognise whether several BrainSites are identical in spite of having different acronyms. For example, in CoCoMac there are four different acronyms for the caudate nucleus: Caud, Cd, C#3, CA#1. When AM is generated, these four acronyms are represented as four different BrainSites of AM. Identity between them is only recognised if at least a certain subset of all these relations have been entered manually by the user. This problem is addressed by the GeneralMap (GM) which is described in section C.2.2.

### **C.2.1.4      *Summary: what are the advantages and disadvantages of introducing AM?***

AM ensures that BrainSites that have identical acronyms but no information on relations between them are automatically linked by I-relations. No special attention of the data collator is required during the data representation process.

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

On the other hand, AM increases computational costs for ORT by increasing the number of BrainSites in the transformation graph.

## **C.2.2 General Map (GM)**

While the co-existence of incongruent parcellation schemes is typical for many brain regions, differences between brain maps are rare in other regions of the brain. This mainly concerns regions with obvious demarcations of structures by macroscopic features (e.g. demarcation by white matter) or because parcellations with very coarse resolution have been considered sufficient so far. In both cases, different authors refer to brain structures without explicit definitions, e.g. they refer to “the” caudate nucleus or “the” superior colliculus. Unfortunately, although the definition of these BrainSites is usually identical between authors (at least at the coarse resolution that they are normally referred to), a variety of different acronyms are used by authors for their designation. For example, four different acronyms from various articles are known for the caudate nucleus in CoCoMac: Caud, Cd, C, CA. Without using GM, the assumed I-relations between BrainSites that use these acronyms to refer to the caudate nucleus have to be entered manually for each pair of BrainSites with different acronyms. This process is not only tedious but also contains the risk that there may be disconnected clusters of identical, yet differently named BrainSites if there are missing relations between these groups of BrainSites.

As a solution to this problem, this section presents the concept of the GeneralMap (GM).

### **C.2.2.1 Definition of the GeneralMap (GM)**

GM is a predefined set of BrainSites that represent general definitions for brain structures. Importantly, GM offers BrainSites for those regions of the brain only where brain structures are commonly referred to (i) without explicit definitions and (ii) by different names in the literature although they are identical (or at least very similar). The BrainSites of GM are defined according to a predefined hierarchy of internal relations. All BrainSites contained by GM

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

as well as the internal hierarchical relations are specified in a separate document. Note that there is no semantic overlap between BrainSites in GM (except as specified by the internal relations). That is, there is only one BrainSite in GM referring to the superior colliculus (GM-SC), but GM-SC has L-relations to each of the seven layers of superior colliculus (GM-SC\_I, ..., GM-SC\_VII).

While all internal relations of GM are predefined, external GM-relations (i.e. relations between elements of GM and BrainSites from original articles) must be specified by the data collators. Practical guidelines for linking original BrainSites to GM are described in section C.2.2.2.

### **C.2.2.2      *Practical guidelines for using GM***

This section describes which BrainSites are at all linked to the GM through relations, when this is necessary, and how the relations are chosen.

**Does every single BrainSite in CoCoMac have to be connected to a corresponding GM-BrainSite?** No. Only those OM-BrainSites in CoCoMac need to be linked to GM that (i) are situated within the brain regions covered by GM and (ii) that have an acronym which is not used by any other OM-BrainSite that has already been linked to GM.

(The second point can be checked quite easily via a query that displays all relations between GM- and OM-BrainSites in CoCoMac. Use the button labelled “List entered relations by map” in the Main Switchboard of CoCoMac and enter “GM” as query parameter. Then search for the GM-BrainSite that would correspond to the OM-BrainSite in question and check whether it has already been linked to another OM-BrainSite with the same acronym.)

In more formal terms: Given an existing BrainSite from BrainMap  $M_1$  with an acronym  $A$  (i.e.  $M_1-A$ ) that is already linked to the corresponding GM-BrainSite (i.e.  $GM-A$ ), another BrainSite from BrainMap  $M_2$  with the same acronym (i.e.  $M_2-A$ ) does not need to be linked explicitly to GM. As long as the concept of the Acronym Map (AM) is used in ORT, the identity of the acronyms ensures that  $M_2-A$  will be linked to both  $M_1-A$  and  $GM-A$  during the optimisation of the transformation graph. **NOTE:** it is the responsibility of the

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

individual data collator to check and decide whether a given OM-BrainSite needs to be linked to GM!

**How are OM-BrainSites linked to GM-BrainSites?** Only I-relations are allowed to link OM-BrainSites to corresponding GM-BrainSites. This means that relations between OM and GM are only allowed for BrainSites at the same conceptual level (e.g. it is not possible to directly link a nucleus in GM to a subnucleus in OM). Seen from another perspective, this also means that, within the brain regions for which it is defined, GM must offer a BrainSite for each conceptual level of data description that is commonly used by the original articles. Overall, the exclusive use of I-relations ensures that OM-BrainSites with different acronyms but identical meaning which are linked to the same GM-BrainSite will themselves be considered to be identical during the optimisation of the transformation graph in ORT.

### **Is GM completely static or can new BrainSites be added to GM?**

GM is neither completely static nor totally flexible. For any given region of the brain, the definition and the hierarchy of the GM-BrainSites must be carefully chosen such that (i) all relevant levels of resolution are covered and yet (ii) finer levels of resolution can still be added should it become necessary in the future. Should any of the data collators feel that the definition of GM is inadequate in any respect, she/he should raise this issue as quickly as possible so that potentially necessary extensions of GM can be discussed.

#### **C.2.2.3      *Summary: what are the advantages and disadvantages of introducing GM?***

GM allows for correct mapping between BrainSites with different acronyms but identical semantics.

However, the correct use of GM also requires some additional efforts by the data collators since it needs to be decided for each new OM-BrainSite whether it needs to be linked to a corresponding GM-BrainSite or not.

### **C.2.3 Comparing AM and GM**

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

The concepts of AM and GM are often best understood when directly contrasted. The following table briefly juxtaposes their main differences.

<b>AcronymMap (AM)</b>	<b>GeneralMap (GM)</b>
AM contains algorithmically created BrainSites. No changes possible for the data collator.	The BrainSites in GM are predefined. Extensions are possible, but must be discussed with the group before.
AM includes one BrainSite for each acronym in CoCoMac that is used by at least two original BrainSites.	GM offers BrainSites for certain regions of the brain only.
The BrainSites of AM have no pre-defined hierarchical relations specified by the data collators.	The BrainSites of GM are defined according to a predefined hierarchy. This hierarchy is specified in a separate document.
In AM, BrainSites can refer to equivalent brain structures (e.g. AM-Caud, AM-Cd, AM-C#3, AM-CA#1 all refer to the caudate nucleus).	In GM, there is no semantic overlap between its elements except for hierarchical relations (e.g. there is only one BrainSite in GM referring to the caudate nucleus as an entity).
Both intrinsic and external relations of AM do not need to be specified by the user. Instead, they are generated algorithmically.	Both internal and external relations of GM must be specified by the user.
AM automatically considers two BrainSites to be identical if they share the same acronym.  AM does not recognise the identical semantics of BrainSites from different articles that refer to the same brain structure but have different acronyms.	OM- BrainSites that refer to the same brain structure but have different acronyms are linked by the user by I-relations to the same GM-BrainSite. For example, BrainSites that are based on the acronyms Caud, Cd, C#3, CA#1 are linked by an I-relation to the GM-BrainSite that represents the caudate nucleus (e.g. GM-NuclCaudatus). The data collator must ensure that the relations between each OM-BrainSite and their counterpart in GM is correct and complete. Relations between the OM-BrainSites will then be constructed by the graph optimisation in ORT.

### C.3 Definition of different types of BrainSites

Each BrainSite in CoCoMac is of a particular type. Currently, 9 different classes of types (*SiteClasses*) are defined that characterize what kind of

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

“module” a given BrainSite represents within the organizational hierarchy of the brain (i.e. whether it is a cortical or subcortical structure, an area, a lamina, a column, a single neuron, etc.). Specifically, each SiteClass contains one or more concrete types of BrainSites (*SiteTypes*). SiteTypes use a suffix (“2D” or “3D”) to indicate whether a given BrainSite is used as a 2D or 3D “storage unit” for the description of connectivity data (i.e. whether any ECs that refer to this BrainSite reflect the extension of information within a 2D or 3D space). Whilst, in principle, there could therefore be two SiteTypes for each SiteClass, this is the case for isocortical areas, allocortical areas, and cortical laminae only. For all other SiteClasses, only 3D representations have been used in the literature contained by CoCoMac so far.

Where in CoCoMac are SiteClasses and SiteTypes used? SiteClasses refer to the nature of BrainSites as such and are thus referred to in the “BrainMaps\_BrainSites” table. SiteTypes are used for the description of experimental results and are thus contained by the tables “Injections”, “Injections\_AffectedSites”, and “LabelledSites\_Data”.

**NOTE:** The main purpose of providing 2D and 3D versions of the SiteType is allow for an exact representation of ECs. Whether a 2D or 3D representation of BrainSites is chosen should depend on the way a given article illustrates its data. That is, if lateral overviews or flat maps are shown, 2D representations are adequate whereas 3D representations better capture the data provided by sections. The distinction between 2D and 3D representations is not absolutely vital as ORT can transfer ECs between 2D and 3D representations of the same BrainSite easily and without much loss of information. At the end of the day, it depends on the data collator what representation he considers to be more appropriate for the data from a given article. As a guideline, however, it is suggested that, if both sections (3D) and overviews / projections (2D) are provided, the latter should be preferred. The reasons for this advice is that (i) 3D representations allow for less discrimination between ECs (e.g. a “complete” (C) extension of label will almost never be found in 3D but relatively often in 2D) and (ii) an EC=C represented in 2D will become EC=X when transformed to 3D (see section on the “Algebra of Transformation”).

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

Also, it is important to remember that sections are generally problematic with regard to their information on Extension Codes (see sections F.3 and G.2 below).

## C.4 Tables that contain mapping data

### C.4.1 BrainMaps\_BrainSites\_SiteClasses

This table contains a single field **SiteClass**. As explained in section C.3, SiteClasses are relatively broad classes for distinguishing different types of BrainSites as envisaged by a modular and hierarchical concept of brain organization. For example, cortical BrainSites can be categorized into cortical and allocortical areas, and each area again can consist of individual laminae, columns, or bands, each of which is made up of neuronal populations whose neurons can be subdivided into different compartments. At the moment, CoCoMac distinguishes between the nine following SiteClasses:

- Area\_IsoCtx
- Area\_AlloCtx
- Lamina\_Ctx
- Band\_Ctx (=intra-areal volume)
- Column\_Ctx
- Nucleus\_SubCtx
- Lamina\_SubCtx
- Neuron\_Ensemble
- Neuron\_Soma

The use of these different SiteClasses are explained in detail in the following paragraph on SiteTypes.

As is easily noted, there is a general distinction between cortical and subcortical BrainSites. This distinction is mainly for computational reasons. By keeping cortical and subcortical data separate throughout most parts of ORT, computational efficiency for searching procedures is considerably increased.

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

**NOTE:** There is one region in the brain where authors disagree whether this should be considered a cortical or subcortical region: the amygdala and the surrounding neural tissue (periamygdaloid region). To maintain consistency, all amygdaloid and periamygdaloid structures are considered to be subcortical BrainSites in CoCoMac.

## C.4.2 BrainMaps\_BrainSites\_SiteTypes

This table contains a single field **SiteType**. This table contains the twelve different “SiteTypes” that CoCoMac currently distinguishes. As explained in section C.2, SiteTypes use a suffix (2D/3D) to indicate whether a given BrainSite is used as a 2D or 3D “storage unit” for the description of connectivity data (i.e. whether any ECs that refer to this BrainSite reflect the extension of information within a 2D or 3D space). Whilst, in principle, there should therefore be two SiteTypes for each SiteClass, this is the case for isocortical areas, allocortical areas, and cortical laminae only. For all other SiteClasses, only 3D representations have been used in the literature contained by CoCoMac so far.

- **Area\_AlloCtx\_2D:** This SiteType is used for the representation of experimental data referring to allocortical BrainSites that are shown by 2D projections, for example, lateral overviews of data projected onto the cortical surface or flat maps (e.g. see the hippocampal brain sites CA1-CA4 in Blatt and Rosene 1998 and Barbas and Blatt 1995).
- **Area\_AlloCtx\_3D:** This SiteType is used for the representation of experimental data referring to allocortical brain sites that are shown within a 3D space (e.g. a labelled area shown by several sections).
- **Area\_IsoCtx\_2D:** This SiteType is used for the representation of experimental data referring to isocortical BrainSites that are shown by 2D projections, for example, lateral overviews of data projected onto the cortical surface or flat maps (e.g. data on visual cortical brain sites shown by Distler et al. 1993).

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

- **Area\_IsoCtx\_3D:** This SiteType is used for the representation of experimental data referring to isocortical BrainSites that are shown within a 3D space (e.g. a labelled area shown by several sections).
- **Band\_Ctx\_3D:** This SiteType is used for the representation of experimental data referring to three-dimensional structurally homogenous bands of tissue across several or all cortical laminae within a cortical area (e.g. thick, thin and pale stripes defined by cytochrome oxidase staining in area V2). In other words, this type refers to "intra-areal volumes". Note: this type is restricted to structurally homogenous neuronal populations in the cortex; for subcortical and highly distributed populations of neurons, choose the type "Neuron\_Ensemble\_3D" instead.
- **Column\_Ctx\_3D:** This SiteType is used for the representation of experimental data referring to cortical columns (conceptualized as 3D elements).
- **Lamina\_Ctx\_2D:** This SiteType is used for the representation of experimental data referring to single cortical laminae that are shown by 2D views, e.g. sections tangential to their main orientation. It should be used for allocortical brain sites that lack the typical six-layered organization of isocortical areas (for example, hippocampal brain sites CA1-CA4 in Yukie 2000) or isocortical brain sites with sublaminae (for example, visual area 17 in Hubel & Wiesel 1972). Each of the laminae in such an area should be represented separately.
- **Lamina\_Ctx\_3D:** This SiteType is used for the representation of experimental data referring to single cortical laminae that are shown by sections perpendicular or oblique to their main orientation. It should be used for allocortical brain sites that lack the typical six-layered organization of isocortical areas (for example, hippocampal brain sites CA1-CA4 in Yukie 2000) or isocortical brain sites with sublaminae (for example, visual area 17 in Hubel & Wiesel 1972). Each of the laminae in such an area should be represented separately.
- **Lamina\_SubCtx\_3D:** This SiteType is used for the representation of experimental data referring to single laminae of subcortical brain sites.

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

These are normally shown by sections. This SiteType should be used for subcortical brain sites with laminar organization (e.g. layers of the superior colliculus in Huerta and Kaas 1988 or the superficial amygdaloid nuclei in Amaral et al. 1992). Each lamina of such layered nuclei should be represented separately.

- **Neuron\_Ensemble\_3D:** This SiteType is used for the representation of experimental data referring to general populations of neurons. These populations can be defined by arbitrary criteria (e.g. all pyramidal cells within a given area, all calbindin-positive cells within a given lamina, all dopaminergic neurons of a given nucleus, etc.). Also, they may consist of spatially distributed neurons or form a continuous block of tissue. This SiteType thus is the most general type above the single neuron level.
- **Neuron\_Soma\_3D:** This SiteType is used for the representation of experimental data referring to the soma of a single neuron.
- **Nucleus\_SubCtx\_3D:** This SiteType is used for the representation of experimental data referring to subcortical nuclei and their subparts. These data are usually represented by sections (e.g. amygdaloid nuclei in Amaral et al. 1992 or thalamic nuclei in Olszewski 1952).

### C.4.3 BrainMaps

- **ID:** The ID of the brain map is identical with the ID of the paper, i.e. there is a 1:1-relation between them.
- **Delineation\_BrainSites:** Tick this box if at least one brain site of the map is delineated (see *Delineation* above).
- **UnspecificAdoption\_BrainSites:** Tick this box if at least one brain site of the brain map is adopted in an ambiguous way (see *Unspecific Adoption* above).
- **Comments** OPTIONAL.
- **dbCollator**

The tables “BrainMaps\_Delineation” and “BrainMaps\_UnspecificAdoption” (see C3.4. and C.3.5) contain further information.

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

### C.4.4 BrainMaps\_Delineation

- **ID\_BrainMaps:** Primary key as well as foreign key, linked to "BrainMaps".
- **Reference\_Text:** Page number(s) where the delineation of brain sites is described.
- **Reference\_Figures:** Figures which provide a general overview of the delineated brain sites (usually not figures used to display the experimental findings!)
- **Region\_Parcellation:** This field describes the region of the cortex that the map refers to (e.g. "prefrontal cortex", "occipital, parietal, and temporal visual areas"). OPTIONAL.
- **ImportantNewBrainSites:** If the article describes a brain site for the first time that subsequently becomes a standard brain site in the literature, this information can be stored in this field. Example: Area LIP was first introduced by AAC85 and has become a standard area since then. OPTIONAL.
- **RelatedMaps:** This field lists other brain maps which are conceptually related to the current map, i.e. brain maps whose brain sites have been re-defined or extended by this brain map. Examples: B09 is almost the same as B05 except for the three areas 32, 35, 43; CP94 mostly subdivides the areas of W40; LMCR93 combines the areas of MLR85 and MLR91 in one map. OPTIONAL.
- **Citations:** Citations with general statements about the delineated brain sites, e.g. explanations of the criteria used for delineation of brain sites, or references to related maps whose brain sites were confirmed, subdivided, changed, or combined. No statements concerning *Unspecific Adoption!* These are to be entered into the next table. OPTIONAL.
- **Comments:** OPTIONAL.
- **dbCollator**

### C.4.5 BrainMaps\_UnspecificAdoption

- **ID\_BrainMaps:** Primary key as well as foreign key, linked to "BrainMaps".

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

- **Reference\_Text:** Page number(s) where the unspecific adoption of brain sites is described.
- **Citations:** Citations which describe Unspecific Adoption are important because they explain why the map of the current article contains brain sites that may appear well defined at first sight. Example: "To describe our findings in the prefrontal cortex, we refer to the nomenclature of standard parcellation schemes (Walker 1940; Carmichael & Price 1994; Preuss & Goldman-Rakic-1991)." OPTIONAL.
- **Comments:** OPTIONAL.
- **dbCollator**

### C.4.6 BrainMaps\_Methods

This table contains further information about the parcellations stored by "BrainMaps". In particular, data on methodological issues of brain site delineation are to be entered here. CAVE: Do not confound this table with the table "Methods" which contains data on tracing methods!

- **ID:** Primary key of this table which is of type AutoNumber, i.e. a number that is automatically assigned to a new entry.
- **ID\_BrainMaps:** ID of the brain map. Foreign key, linked to "BrainMaps".
- **Criterion:** Methodological criterion that was used to delineate brain sites in the current map (e.g. cytoarchitecture, myeloarchitecture, etc.). Foreign key, linked to "BrainMaps\_Criteria".
- **MethodologicalDetails:** Adds further explanation to the criterion above, e.g. what specific type of staining was used (Nissl, Gallyas, SMI32...).
- **Reference\_Text:** Page number(s) where the author describes the methods for delineation of brain sites.
- **dbCollator**

### C.4.7 BrainMaps\_BrainSiteAcronyms

This table contains all acronyms that are used for the designation of brain sites. If there are abbreviations, which are made up of the same letters but

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

designate different brain sites, add an “#” and an unambiguous number (e.g. “Ig” is used to designate both “Granular insula” and “Indusium griseum”, therefore these two brain sites are referred to by “Ig#1” and “Ig#2”). Note that Access does not distinguish between uppercase and lowercase letters - i.e. “Ig” and “IG” are treated as identical strings!

**NOTE:** As described in the Administration Guidelines, make sure to disable **all** options in the Tools/AutoCorrect menu of Access, otherwise Access might autocorrect certain Brainsite Acronyms, e.g. “VAmc” is corrected to “Vamc”, which does not correspond to the original nomenclature of the article.

- **Acronym:** Abbreviation of the full name of a brain site.
- **FullName:** Explanation of the acronym.
- **dbCollator**

### C.4.8 BrainMaps\_BrainSites

For each article, this table lists all BrainSites that were either explicitly defined (*Delineation*) or unspecifically adopted from another map (*Unspecific Adoption*). Within CoCoMac, this set of BrainSites is called the “BrainMap” or briefly “map” of the article. For each BrainSites of each BrainMap, a unique ID is created by joining the ID of the BrainMap and the acronym of the BrainSite via a dash (e.g. area 6 of Brodmann's map of 1905 is referred to by “B05-6”). If the acronym from “BrainMaps\_BrainSiteAcronyms” contains a number sign (#), this should be omitted when creating the ID. **NOTE:** This ID is essential for all tables that refer to BrainSites in any way!

The following rules are important to follow:

Brain sites referred to by the author **are not added** to this table if

- Specific Adoption: The author specifies clearly which other parcellation schemes are referred to (e.g. “Label was found in Brodmann's (1905) area 6 and von Bonin and Bailey's (1947) area FCBm.”). In this case, the areas of the referred parcellation scheme are used to describe the experimental data.
- Topographic description: The article does not refer to a delineated brain site by its name or acronym, but (i) uses broad topographic labels to

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

describe its data OR (ii) displays data on figures which lack areal names and borders and for which none of the map(s) defined or adopted by the article is valid. Example: "Label was also found in the vicinity of the principal sulcus."

In this case, a brain site defined by another map is used to describe the experimental data.

**NOTE:** The PDC\_Site must be P for this case!

The names of the brain sites as used by the author **are added** to this table only in one of the following cases:

- Delineation: The article deals with the investigation of areal borders and either presents a new parcellation scheme or modifies/confirms an existing one. The locations and boundaries of the brain sites may be described in the text or be merely displayed by general figures of brain maps. The field "SiteDef\_Type" is marked as D (delineated).
- Unspecific Adoption – case 1: The article uses a parcellation scheme which is a mixture of several other maps and which makes it difficult to decide for a given brain site what scheme it was taken from. The field "SiteDef\_Type" is marked as A (adopted).
- Unspecific Adoption – case 2: The author uses a term for a brain site without specifying the underlying parcellation scheme directly (e.g. referring to an "area 7" without stating which area 7 is meant). The field "SiteDef\_Type" is marked as A (adopted).

**NOTE:** As described in the Administration Guidelines, make sure to disable **all** options in the Tools/AutoCorrect menu of Access, otherwise Access might autocorrect certain Brainsite Acronyms, e.g. "VAmc" is corrected to "Vamc", which does not correspond to the original nomenclature of the article.

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

Data fields of "BrainMaps\_BrainSites":

- **ID**: ID for a particular BrainSite of a particular BrainMap (construction rules – see above). Primary key.
- **ID\_Acronyms\_BrainSites**: Acronym of the BrainSite. Foreign key, linked to "BrainMaps\_BrainSiteAcronyms".
- **ID\_BrainMap**: ID of the BrainMap. Foreign key, linked to "BrainMaps".
- **SiteDef\_Type**: States whether the brain site has been delineated (D) or unspecifically adopted (A) -see above.
- **SiteClass**: Then SiteClass of this BrainSite (see section C.2)
- **dbCollator**
- **SourceData**: Future option to determine which brain sites are accepted as input for the mapping by ORT. Not implement yet. Will only be of importance for the ORT algorithms.
- **TargetMap**: Tick box that allows to define which brain sites are included in the target map for a transformation by ORT. Only of importance for the ORT algorithms.

## C.4.9 InterMapRelations

This table stores the logical relations (mappings) between BrainSites in different BrainMaps in the form RC / Site A / Site B. This format is to be read as follows:

I / A / B: A is *identical* to B.

S / A / B: A is a *subarea* of B.

L / A / B: A is "*larger*" than B (i.e. B is a subarea of A).

O / A / B: A and B *overlap*.

C / A / B: A is a lamina of B (more exactly: 2D or 3D lamina A is *collapsed* onto 2D area B)

E / A / B: A includes lamina B (more exactly: 2D area A is *expanded* to 2D or 3D lamina B)

Data fields of this table:

- **RelationCode**: I, S, L, O, C, or E. Part of primary key.

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

- **BrainSite\_A**: Foreign key and part of primary key, linked to "BrainMaps\_BrainSites".
- **BrainSite\_B**: Foreign key and part of primary key, linked to "BrainMaps\_BrainSites".
- **dbCollator**

There are fixed rules which combinations of Relation Codes (RCs) and SiteClasses of BrainSites within InterMapRelations are allowed. All permissible combinations are summarized by the following table. To ensure logical consistency of the relation data, the CoCoMac data entry interface will check the BrainSites within each entered relation and reject any RC that is not listed by the table below. Red entries denote constellations of RCs and SiteClasses for which ORT may have to perform a conversion between 2D/3D objects when transforming data to a new parcellation scheme (e.g. when mapping data from a 3D cortical column to a 2D cortical area). This does not affect the data entry process at all – the user simply has to follow the table below.

	Area_IsoCtx	Area_AlloCtx	Lamina_Ctx	Column_Ctx	Band_Ctx	Nucl_SubCtx	Lamina_SubCtx	Neuron_Soma	Neuron_Ensemble
Area_IsoCtx	ISLO	ISLO	E	L	L	-	-	L	ISLO
Area_AlloCtx	ISLO	ISLO	E	L	L	-	-	L	ISLO
Lamina_Ctx	C	C	ISLO	O	O	-	-	L	ISLO
Column_Ctx	S	S	O	ISLO	SO	-	-	L	ISLO
Band_Ctx	S	S	O	LO	ISLO	-	-	L	ISLO
Nucl_SubCtx	-	-	-	-	-	ISLO	L	L	ISLO
Lamina_SubCtx	-	-	-	-	-	S	ISLO	L	ISLO
Neuron_Soma	S	S	S	S	S	S	S	I	S
Neuron_Ensemble	ISLO	ISLO	ISLO	ISLO	ISLO	ISLO	ISLO	L	ISLO

Permissible constellations of RCs and SiteClasses within InterMapRelations.

**NOTE:** Never define a relation across more than one level. For example, suppose that you deal with a cortical area X for which subareas  $Y_1 \dots Y_n$  are

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

described by the same map. Suppose further that, within that map, area  $Y_i$  is further subdivided into several areas  $Z_1 \dots Z_m$ . Although possible from a logical point of view, you should not try to define relations like  $Z_1 -S \rightarrow X$ . Instead, the correct way would be to define relations that span only one level each, for example  $Z_1 -S \rightarrow Y_i$  and  $Y_i -S \rightarrow X$ .

The reason for this constraint is that whenever ORT evaluates an S- or O-relation, it needs to determine how many areas are, in total, equivalent to or include the target site. It does so by finding all other areas that have an S- or O-relation to the target site. If we had defined both  $Z_1 -S \rightarrow X$  and  $Y_i -S \rightarrow X$ , ORT would falsely assume that  $Z_1$  and  $Y$  would be independent sites, both of which would have to be searched for experimental data. This is, of course, wrong since  $Z_1$  is included by  $Y$  and thus only  $Y$  needs to be considered by ORT.

**NOTE:** Although the order in which two BrainSites are included into a relation does not affect the relation as such (e.g. S/X/Y represents the same relation as L/Y/X) we nevertheless need a clear rule to avoid multiple representations of the same relation within CoCoMac. Therefore, the following guideline should be applied: Assign the more recently defined BrainSite to the A-position of the relation. If both BrainSites were defined in the same year, assign the brain site with the alphabetically lower name to the A-position of the relation. Examples: Dealing with the BrainSites B05-6 and BB47-FBA, the relation should be expressed as S / BB47-FBA / B05-6 (and *not* as the semantically equivalent relation L / B05-6 / BB47-FBA). Dealing with the identical BrainSites BDU91-V2 and FV91-V2, the relation should be expressed as I / BDU91-V2 / FV91-V2 (and *not* as the semantically equivalent I / FV91-V2 / BDU91-V2).

Don't despair if this sounds complicated: all entries are controlled for compatibility with this scheme by an algorithm running in the background of the form by which you enter the InterMapRelations. If something is not quite right, it will be corrected automatically for you.

### C.4.10 InterMapRelations\_References

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

This table stores the references on which the relations between brain maps are based and also stores the associated PD codes.

- **RelationCode:** Part of primary key and foreign key linked to "InterMapRelations".
- **BrainSite\_A:** Part of primary key and foreign key linked to "InterMapRelations".
- **BrainSite\_B:** Part of primary key and foreign key linked to "InterMapRelations".
- **ID\_Literature:** ID of the paper where the statement about the relation is given. Part of primary key and foreign key linked to "Literature".
- **Reference\_Text:** Page number(s) where the statement about the relation is given. If there is no textual information about this relation, enter a dash ("-"). Part of primary key.
- **Reference\_Figures:** Figure number(s) where the statement about the relation is given. If there are no figures with information about this relation, enter a dash ("-"). Part of primary key.
- **PDC\_Relation:** PD code of the relation, linked to "PDC\_Relation".
- **Citation:** Citation of the statement, which describes the relation.  
OPTIONAL - but it should be included whenever possible!!! Add the page number in brackets after the citation – this is especially important if there are several relevant statements and the field "Reference: Text" thus refers to several pages.
- **Comments:** Whenever the relation is not stated unambiguously, you should try to briefly describe your reasons for assigning a specific PD code. This especially applies if any further references are taken into account (i.e. PD codes F+G, J+K, N+O) or if you are forced to make an assumption (PD code P).
- **dbCollator**

## ***D. Experimental data***

### **D.1 Tables containing experimental data**

#### **D.1.1 Methods**

This table contains information about the methods used for the tracing experiments. **NOTE:** Do not confound this table with the table "BrainMaps\_Methods" which contains data on the methods used for delineating brain sites within a given brain map!

- **ID:** Primary key of this table, which is of type AutoNumber i.e. a number that is automatically assigned to a new entry.
- **ID\_Literature:** ID of the paper. Foreign key, linked to "Literature".
- **Reference\_Text:** Page number(s) where the tracing methods are described.
- **Reference\_Figures:** Figure number(s) where the statement about the method is given. If there are no figures with information about the method, enter a dash ("-").
- **TracerSubstance:** Used tracer substance. Foreign key, linked to "Methods\_TracerSubstances".
- **BilateralUse:** This box is ticked if the tracer substance was used for injections in both hemispheres within the same animal – regardless whether the same or different brain structures were injected.  
**NOTE:** If bilateral injections with the same tracer are applied to a subset of the animals only, you need to define two methods: one with a ticked "Bilateral Use"-field, one without.
- **InjectionMethod:** Specifies whether a pressure injection (P), a iontophoretic injection (I), or diffusion (D) (e.g. from HRP gel implants) was used to apply the tracer substance. If unknown, enter a question tag (?).

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

- **SurvivalTime:** Time between the injection and the death of the animals in hours (h) or days (d).
- **Thickness\_Sections:** Thickness of the histological sections in  $\mu\text{m}$  (micrometer).
- **Comments:** OPTIONAL
- **dbCollator**

Some articles written by the same authors refer to the same set of injections that are described only in one of the articles (e.g. CP95b and CP96 refer to the methods and injections described by CP95a without giving a detailed description themselves). In this case, the injection methods and injections themselves are described explicitly for each of the concerned articles, even in those articles that do not describe methods or injections but only refer to a companion article. These “virtual” representation of methods / injections avoid confusion and allow to algorithmically relate experimental data to their correct article, Still the “Reference\_Text” and “Reference\_Figures” fields should point to the references in the original article.

### D.1.2 Methods\_Animals

- **ID\_Method:** ID of the method. Primary and foreign key, linked to "Methods".
- **Nbr\_AllAnimals:** Number of animals across all species used for tracing experiments with the associated method.
- **Nbr\_AllFemaleAnimals:** Number of all female animals across all species (number of all male animals can be calculated by subtracting Nbr\_AllFemaleAnimals from Nbr\_AllAnimals). OPTIONAL.
- **AgeRange:** OPTIONAL. If unknown, enter a question mark (?).
- **WeightRange:** OPTIONAL. If unknown, enter a question mark (?).
- **Comments:** OPTIONAL.
- **dbCollator**

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

### D.1.3 Methods\_Animals\_Details

- **ID\_Methods\_Animals:** ID of the associated entry in "Methods\_Animals". Part of primary key. Foreign key, linked to "Methods\_Animals".
- **MacaqueSpecies:** Species used for the tracing experiments (e.g. *Macaca mulatta* [= Rhesus], *Macaca fascicularis* [= *Macaca cynomolgus*], *Macaca nemestrina* etc.). Part of primary key.
- **Gender:** Sex of the animals (F = female, M = male, ? = unknown). Part of primary key.
- **Nbr:** Number of animals of the species and the gender indicated by the "MacaqueSpecies" and "Gender" fields. If unknown, enter a question mark (?).
- **dbCollator**

**NOTE:** The difference between "Methods\_Animals" and "Methods\_Animals\_Details" is that the former is related to all animals used for a particular tracing method, whereas the latter indicates the species and gender of homogeneous subsets of animals.

### D.1.4 Methods\_RadioactiveTracers

This table contains data on radioactive tracer substances used for the tracing experiments.

- **ID\_Method:** Foreign key, linked to "Methods". Part of primary key.
- **TracerSubstance:** Radioactive tracer substance (e.g. [3H]-proline). Part of primary key.
- **Activity:** Radioactive activity in  $\mu\text{Ci}$  /  $\mu\text{l}$ . OPTIONAL.
- **SpecificActivity:** Specific activity in Ci / mmol. OPTIONAL.
- **Ratio:** Ratio between this substance used and all other radioactive substances used in the injection (in %). Example: MBG91 used a 50%-50% mixture of [3H]-proline and [3H]-leucine for their injections. OPTIONAL.
- **dbCollator**

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

## D.1.5 Injections

This table stores information about the individual injections. If more than one BrainSite is affected by the injected tracer substance, choose the BrainSite which is most strongly affected as "ID\_BrainSite" in this table. If there is no clear information about this in the text, try to evaluate figures. If they do not provide clear information either, choose the first of the affected BrainSites named by the text. List the other BrainSites in the table

"Injections\_AffectedSites" (see below).

- **ID:** Primary key which has the following format: <ID\_Literature>-<description of injection>. Depending on the article, the latter term <description of injection> can consist of a case number, the name of a BrainSite, a tracer substance, or the number of a figure - whatever allows a clear identification of the injection. Some examples: BG93-c1...BG93-c13, BDU91-NY, BP87-HRP-6DC.
- **ID\_Method:** ID of tracing method that was used for this injection. Foreign key, linked to "Methods".
- **ID\_BrainSite:** Name of the injection site. Foreign key, linked to "BrainMaps\_BrainSites".
- **PDC\_Site:** PD code of the identification of the injected BrainSite. Foreign key, linked to "PDC\_BrainSiteIdentification".
- **SiteType:** SiteType of the injected BrainSite (see C.2). Foreign key, linked to "BrainMaps\_BrainSites\_SiteTypes".
- **Citation:** Citation of the author's description of the injection and (if present) methodical problems.
- **Reference\_Text:** Page number(s) where the injection is described.
- **Reference\_Figures:** Figures that show the injection.
- **Hemisphere:** States whether the injection was placed in the left (L), right (R), or an unknown (?) hemisphere.
- **EC:** Extension Code of the injection. See general information on Extension Codes below.
- **PDC\_EC:** PD code of the EC of the injection. Foreign key, linked to "PDC\_EC".

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

- **Volume:** Volume of injected tracer substance in microliters ( $\mu\text{l}$ ). If unknown, enter a question mark (?).
- **Concentration:** Concentration of injected tracer substance in percent (%). If unknown, enter a question mark (?).  
**NOTE:** this field is not intended to store activity values for radioactive substances – use table Methods\_RadioactiveTracers for this purpose (see D.1.4).
- **AffectedNeighbours:** Tick this box if the injected tracer substance has spread to neighbouring BrainSites. Information about these BrainSites is stored by "Injections\_AffectedSites".
- **MethodicalProblems:** Tick this box if there are methodical problems that make the results of this injection difficult to interpret. Examples: (i) spread of the injection to the white matter, (ii) not all affected sites are being described ("TAA was injected in the prefrontal granular cortex (including area 46)...", p. 31 in MM82b), (iii) surgical damage to the injected brain site or the underlying white matter. Please take care to cite the relevant description in the Citation field above.
- **Comments:** OPTIONAL.
- **dbCollator**

**NOTE:** Some articles written by the same authors refer to the same set of injections that are described only in one of the articles (e.g. CP95b and CP96 refer to the methods and injections described by CP95a without giving a detailed description themselves). In this case, the injection methods and injections themselves are described explicitly for each of the concerned articles, even in those articles that do not describe methods or injections but only refer to a companion article. These “virtual” representation of methods / injections avoid confusion and allow to algorithmically relate experimental data to their correct article, Still the “Reference\_Text” and “Reference\_Figures” fields should point to the references in the original article.

### D.1.6 Injections\_AffectedSites

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

This table lists all BrainSites that were affected by an injection in addition to the main target BrainSite - only if there were any, of course.

- **ID\_Injection:** ID of the injection. Foreign key, linked to "Injections". Part of primary key.
- **ID\_AffectedSite:** ID of the BrainSite that was affected by the injection. Foreign key, linked to "BrainMaps\_BrainSites". Part of primary key.
- **PDC\_Site:** PD code of the identification of the affected BrainSite. Foreign key, linked to "PDC\_BrainSiteIdentification".
- **SiteType:** SiteType of the affected BrainSite (see C.2). Foreign key, linked to " BrainMaps\_BrainSites\_SpecSiteTypes ".
- **EC:** Extension Code of the injection. See general information on Extension Codes below.
- **PDC\_EC:** PD code of the EC of the injection. Foreign key, linked to "PDC\_EC".
- **dbCollator**

### D.1.7 Injections\_Laminae

This table stores information about the laminae that were infiltrated by the injection. An entry is only necessary if the article delivers relevant information (i.e. you do not need to enter "?????" for an unknown laminar pattern of infiltrated tracer substance).

- **ID\_Injection:** ID of the injection. Foreign key, linked to "Injections". Part of primary key.
- **Laminae:** Code for the laminae that were infiltrated by the injected tracer substance. See general information on representation of laminar patterns below.
- **PDC\_Laminae:** PD code of the laminar pattern. Foreign key, linked to "PDC\_Laminae".
- **dbCollator**

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

**NOTE:** Since laminae of isocortical visual area V1, allocortical or subcortical brain sites are entered as distinct representation objects, this table does not need to be considered for their representation.

## D.1.8 LabelledSites\_Descriptions

This table stores basic information that is equally valid for a set of experimental results resulting from one injection (e.g. ipsi- or contralateral labelling, references, etc.). Thus, it prevents that this information is repeated for each single labelled BrainSite described by "LabelledSites\_Data" (see below).

- **ID:** Primary key of this table which is of type AutoNumber, i.e. a number that is automatically assigned to a new entry.
- **ID\_Injection:** ID of the injection from which the labelling resulted. Foreign key, linked to "Injections".
- **Citation:** Author's description of the labelling. OPTIONAL.
- **AddedDescription:** Additional description by the dbCollator of anything that is worth noticing (e.g. particular patterns, like strip- or blob-wise patterns). Example: See entries for BDU91. OPTIONAL.
- **Reference\_Text:** Page number(s) where the statement about the relation is given.
- **Reference\_Figures:** Figure(s) which show the labelling.
- **Hemisphere:** Ipsi- (I) or contralateral (C) labelling with respect to the injected hemisphere. If unknown, enter a question mark (?)
- **Terminal\_vs\_Soma:** Labelling of terminals (T) or somata (S). Analyse carefully those reports which use HRP-WGA that is transported both antero- and retrogradely.
- **Comments:** OPTIONAL.
- **dbCollator**

## D.1.9 LabelledSites\_Data

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

This table describes the actual experimental findings of tracing experiments, i.e. the labelling (or non-labelling) of individual BrainSite.

- **ID:** Primary key of this table which is of type AutoNumber, i.e. a number that is automatically assigned to a new entry.
- **ID\_Description:** ID of the description of the labelled BrainSite. Foreign key, linked to "LabelledSites\_Descriptions".
- **ID\_BrainSite:** ID of the (un)labelled brain site. Foreign key, linked to "BrainMaps\_BrainSites".
- **SiteType:** Specific SiteType of the labelled BrainSite (see C.2). Foreign key, linked to "BrainMaps\_BrainSites\_SpecSiteTypes".
- **PDC\_Site:** PD code of the identification of the (un)labelled BrainSite. Foreign key, linked to "PDC\_BrainSiteIdentification".
- **EC:** Extension Code of the (un)labelled BrainSite. See general information on Extension Codes below.
- **PDC\_EC:** PD code of the EC of the injection. Foreign key, linked to "PDC\_EC".
- **Density:** Density of the labelling. See general information on coding density below.
- **PDC\_Density:** PD code of the density of the label. Foreign key, linked to "PDC\_Density".
- **QuantitativeData:** Tick this box if there are any quantitative data about labelled neurons (e.g. total number of labelled neurons per brain site). The data is then stored in the respective table for quantitative data (one of the "QD" tables, see below).
- **Comments: OPTIONAL**
- **dbCollator**
- **Inconsistencies:** This field is ticked automatically if a control algorithm detects any inconsistencies in this entry (e.g. an unlabelled area that has a density higher than zero). Check carefully any entries with this box ticked.

### D.1.10 LabelledSites\_Data\_Laminae

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

This table stores information about the laminar pattern of labelling for a particular BrainSite in a particular tracing experiment. An entry is only necessary if the article delivers relevant information (i.e. you do not have to enter "?????" for an unknown laminar pattern of labelling).

- **ID\_LabelledSite\_Data:** ID of the entry which contains the data on the labelled BrainSite. Foreign key, linked to "LabelledSite\_Data". Part of primary key.
- **Laminae:** Code for the laminar pattern of labelled cells or terminals. See general information on representation of laminar patterns below. Part of primary key.
- **PDC\_Laminae:** PD code of the laminar pattern. Foreign key, linked to "PDC\_Laminae". Part of primary key.
- **Sublaminae:** Tick this box if there is information on sublaminae. The respective information has to be stored in the table "LabelledSites\_Data\_Sublaminae". Example: See entries of RV94.
- **dbCollator**

**NOTE:** It is not uncommon that a particular brain site (especially if it is a larger one) shows several different laminar patterns of labelling. These should all be stored in this table as independent entries.

**NOTE:** Since laminae of isocortical visual area V1, allocortical or subcortical brain sites are entered as distinct representation objects, this table is not relevant for their representation.

### D.1.11 LabelledSites\_Data\_Sublaminae

This table contains information about labelling of sublaminae as identified by the respective authors.

- **ID\_LabelledSite\_Data:** Foreign key, linked to "LabelledSites\_Data". Part of primary key.
- **Sublamina:** Sublamina that was identified as labelled. Part of primary key.
- **dbCollator**

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

**NOTE:** Since sublaminae of isocortical visual area V1 are entered as distinct representation objects, this table is not relevant for their representation.

### D.1.12 LabelledSites\_Data\_QD\_Site\_Nbr

This table contains quantitative data, namely information about the number of labelled neurons within one BrainSite.

- **ID\_LabelledSite\_Data:** Foreign key, linked to "LabelledSites\_Data".  
Primary key.
- **TotalNbr\_Neurons:** Total number of labelled neurons in one BrainSite.
- **dbCollator**

### D.1.13 LabelledSites\_Data\_QD\_Site\_Percent

This table contains information about quantitative data, namely information about the percentage of labelled neurons within one BrainSite with regard to the total number of labelled neurons resulting from a particular injection.

- **ID\_LabelledSite\_Data:** Foreign key, linked to "LabelledSites\_Data".  
Primary key.
- **Percentage:** Proportion of labelled neurons in one BrainSite to all labelled neurons
- **dbCollator**

### D.1.14 LabelledSites\_Data\_QD\_Lamina\_Percent

This table contains information about quantitative data, namely information about the percentage of labelled neurons within a particular lamina of a particular BrainSite with regard to the total number of labelled neurons within that BrainSite after a particular injection.

- **ID\_LabelledSite\_Data:** Foreign key, linked to "LabelledSites\_Data".  
Primary key.
- **Lamina:** Labelled lamina that the percentage value refers to

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

- **Percentage:** Proportion between labelled neurons in this lamina and labelled neurons in all laminae of this BrainSite.
- **dbCollator**

**NOTE:** Since laminae of isocortical visual area V1, allocortical or subcortical brain sites are entered as distinct representation objects, this table is not used for these BrainSites. Instead, enter the percentage value in LabelledSites\_Data\_QD\_Site\_Percent.

## ***E. Other tables***

### **E.1 Tables only used by ORT algorithms**

- ORT\_AM\_BrainSites\_Ctx
- ORT\_AM\_BrainSites\_Subctx
- ORT\_BrainSites\_Ctx
- ORT\_BrainSites\_Subctx
- ORT\_IntegratedPrimaryProjections
- ORT\_IntegratedRelations
- ORT\_IntegratedResultingProjections
- ORT\_OptimizedTransformationGraph\_Ctx
- ORT\_OptimizedTransformationGraph\_Subctx
- ORT\_PrimaryProjections
- ORT\_ResultingProjections
- ORT\_tmpTable\_AreasInCurrentPath
- ORT\_Tracking\_ExpData\_1 ... ORT\_Tracking\_ExpData\_5

### **E.2 Base tables (i.e. without foreign keys)**

- BrainMaps\_BrainSiteAcronyms
- BrainMaps\_BrainSites\_SiteClasses
- BrainMaps\_BrainSites\_SiteTypes
- BrainMaps\_Criteria
- dbCollators
- Literature
- Literature\_Abbreviations\_Journals
- Literature\_Authors
- Methods\_TracerSubstances
- PDC\_BrainSiteIdentification
- PDC\_EC

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

- PDC\_Density
- PDC\_Laminae
- PDC\_Relations
- StatusOptions\_DataEntry
- StatusOptions\_Proofreading

## E.3 Administration tables

- Administration\_DataCount
- Administration\_LogTable
- Administration\_Proofreading\_TracingData
- Administration\_Proofreading\_MappingData
- Administration\_RemainingTasks\_dbStructure
- Administration\_RemainingTasks\_Papers

## ***F. General coding rules***

### **F.1 Choosing appropriate types of BrainSites**

Choosing the virtual “spatial dimensionality” of the SiteTypes for the description of experimental data should be guided by the practicalities of representing data in CoCoMac. That is, if the article mainly uses overviews to present a certain experimental finding, 2D-representation objects should be chosen (e.g. Area\_IsoCtx\_2D, Area\_AlloCtx\_2D). On the other hand, if the article mainly uses sections to present a certain experimental finding, 3D-representation objects are preferable (e.g. Area\_IsoCtx\_3D, Area\_AlloCtx\_3D). However, both forms of data representation may co-exist within a given article therefore it is ultimately the collator's choice what dimensionality of representation he chooses for a particular injected or labelled BrainSite. As a guideline, however, it is suggested that, if both sections (3D) and overviews / projections (2D) are provided, the latter should be preferred. The main reason for this advice is that sections are generally problematic for assessing the extent of label or injected tracer substance within a given brain site (see section F.3).

**NOTE:** Please note that the SiteType of a given BrainSite is only important for correct interpretation of the EC (i.e. whether an EC is to be interpreted as label in 2D or 3D space), but not for the identification of the BrainSite as such, i.e. the PDC\_Site is independent of the chosen SiteType! For example, if a particular article provides both sections and overviews concerning a particular BrainSite, then it is perfectly possible to use information provided by the sections (e.g. names and/or borders of BrainSites) for determining the PDC\_Site and to use the overview figure for determining the EC and PDC\_EC. In other words, if an overview figure does not provide precise information on names and/or borders of brain sites, than a single section showing label for a brain site identified by name and borders is absolutely sufficient for securely identifying this brain site to be labelled. Thus the

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

definition of the PDC\_Site can be based on information from this section, whereas the PDC\_EC is more easily defined by evaluating the information provided by overview figures even though the latter may not show precise borders. On the contrary, even several non-labelled sections are not sufficient to identify a given brain site as unlabelled because identification of non-labelled brain sites is only possible from figures that show the entire extent of the brain site (i.e. that is overview figures). In this case, definition of the PDC\_Site is best based on information from overview figures even if they lack precise names and/or borders (note that the PDC\_EC generally is "-" [i.e. dash] for unlabelled brain sites).

## F.2 Identification of (un)labelled brain sites

When evaluating the results of a particular injection, proceed in the following manner:

First, read the text and/or table in which the results are described textually.

For each brain site mentioned by the text to be (un)labelled, evaluate the figures whether they confirm or contradict this statement. If there are mismatches and contradictions, the information as given by the text should be preferred. (**NOTE:** It is not necessarily a contradiction if a brain site is stated by the text to be labelled, and the figures show *sections* through this brain site without label. The sections might miss the labelled part of the brain site).

Second, evaluate the figures for additional results that were not described by the text:

- (a) Brain sites that the text does not mention explicitly as labelled, but are shown labelled by the figures or tables, are always entered as "labelled".
- (b) Brain sites that the text does not mention explicitly as unlabelled, but are shown unlabelled by the figures or tables are only included as "unlabelled" if
  - the respective brain sites belong to those parts of the brain the report generally deals with. Use the same parcellation scheme that the author used for this region.

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

- the article explicitly states that its figures reflect the labelling as found throughout the whole brain/cortex.

In both cases, if the respective brain sites do not belong to the part of the brain for which the article presents or references a parcellation scheme, you need to interpret the location of the (missing) label on the basis of another, self-chosen map. In order to minimize interpretation, use classical parcellation schemes with big areas like B09 or BB47 (or W40 for prefrontal cortex) and apply the appropriate PD code (PDC\_Site=P). Try to determine the location of a particular brain site by its relative position to macroscopical landmarks (e.g. sulci). If in spite of all attempts you remain uncertain whether this brain site is labelled or not (e.g. due to lack of areal borders and uncertainty about relative position to landmarks), then refrain from entering any data about this brain site and put a remark into the "Comments" field of "LabelledSites\_Descriptions".

After you have identified the brain sites that will be entered as labelled or unlabelled, the next step is to determine the Extension Codes for the labelled brain sites (by definition, the unlabelled brain sites are all assigned EC = N).

### F.3 Extension Codes (ECs)

There are 4 different Extension Codes that can be used to generally describe the spatial extent of information in a particular brain site A:

**EC (A) = N:** The information is valid for **No** part of A.

**EC (A) = P:** The information is valid **Partially** for A, i.e. there are subparts for which it is not valid.

**EC (A) = C:** The information is valid for the **Complete** extent of A, i.e. for every subpart of A.

**EC (A) = X:** The information **eXists** for A, i.e. due to lack of precise information it is valid at least for a part of A, maybe even for its complete extent.

In our context, this information consists of the spatial extent of injected or transported tracer substance.

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

As with the identification of (un)labelled brain sites, first evaluate the text for any information. Pay attention to typical statements like "label was confined to the dorsal part of brain site X" or "labelled neurons were found throughout the whole extent of brain site Y". Check the figures for confirmation or contradiction of these statements. Then, for all labelled brain sites without textual information on the EC, evaluate the figures:

- If explicit areal borders are shown in the figures, it is relatively easy to extract information about the extension of the label within the brain sites.
- If clear areal borders lack, then EC=P should only be chosen under the condition that the label covers definitely not more than a subpart of the brain site of interest, e.g. if only a small focus of label is present in an brain site of large size. Similarly, EC=C should only be chosen under the condition that the label definitely covers a region that includes the brain site of interest. If none of these conditions are fulfilled, EC=X must be chosen.
- With respect to information on the EC, figures showing brain sections are problematic as they only represent a tiny fraction of an brain site in question, e.g. even in an brain site that is completely filled with label at small regular intervals, some sections may show missing or incomplete labeling. On the other hand, within a partially labelled brain site a given section might still show label throughout its entire extent.
  - ⇒ **Therefore, information provided by sections is not suitable for confirming or refuting textual information on the extension of label** (i.e.  $PDC\_EC \notin \{A, B, D\}$  for this case). **Only if no textual information on the EC is given at all, information provided by sections may be used** (i.e.  $PDC\_EC \in \{L, M, N, O\}$ ).
- If the figures consist of both overviews and sections, the information of the overviews is more authoritative (see above).

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

## F.4 Density coding

"Density of label" and "strength of label" are relatively badly defined concepts that are nevertheless used by many tracing articles in an attempt to grade the strengths of different projections. Although there is no definition universally agreed upon and the terms are usually used somewhat interchangeably, there seems to be a subtle difference between "strength" and "density". "Strength" may be defined by the total number of labelled neurons in a particular brain site, whereas "density" refers to the spatial concentration of labelled neurons (i.e. neurons per brain site unit - see CP95b, p. 644, as an example). Thus, one can envisage a strong connection with low density (i.e. a large number of labelled neurons distributed regularly, without high concentrations over a large brain site) as well as a dense projection with low strength (i.e. few labelled neurons that constitute a small, but highly packed cluster). Most articles are not specific about these concepts and use "strong connections" and "dense connections" interchangeably. In any case, the grading is rather coarse, and we thus choose a pragmatic approach towards these data by including both reports on strength and density.

Textual information about density/strength should contain one of the following keywords:

"weak" / "sparse" / "light" labelling	=> density = 1
"moderate" / "medium" labelling	=> density = 2
"strong" / "dense" / "heavy" labelling	=> density = 3

Information provided by figures is only unambiguous if there are symbols that clearly assign the label within a brain site to one of the three classes above (e.g. if triangles denote weak, circles moderate, and squares strong labelling). Figures that indirectly display the density of label by plotting individual labelled neurons can be evaluated, but the information is certainly not unambiguous, and the appropriate PD code (i.e. PDC\_Density = I if only figures give information on density, or PDC\_Density= C / E if textual information on density exists as well) must be assigned to indicate the strong degree of interpretation. Note that interpreting figures in this way is only possible if it is

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

clearly stated that each dot in the figure represents one or several labelled neurons.

If label of different density/strength is found within the same brain site, the maximal degree of labelling is represented.

## F.5 Laminar patterns

The following principle is applied to represent data on laminar patterns: For each lamina, one alphanumeric placeholder indicates whether label or not is present and the relative strength of label. Thus, laminar patterns in isocortical brain sites are represented by a 6-characters code. Placeholders are:

- **0**: Lamina contains no label.
- **1**: Lamina is weakly / sparsely labelled in comparison to other laminae of the same brain site.
- **2**: Lamina is moderately labelled in comparison to other laminae of the same brain site.
- **3**: Lamina is densely / heavily labelled in comparison to other laminae of the same brain site.
- **X**: Lamina is labelled, however, there is no data on the relative density in comparison to other laminae.
- **?**: The data about this lamina is not clear, therefore it remains unclear whether it contains label or not.

**NOTE:** Since laminae of isocortical visual area V1, allocortical or subcortical brain sites are entered as distinct representation objects, this coding does not apply to them.

**NOTE:** The PDC\_Laminae only concern the laminar pattern as such (i.e. existing vs. non-existing for each individual lamina - the "binary" pattern), but not the relative strengths of the label!

## F.6 Special considerations

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

Some constellations merit special attention. First, in the case of an unlabelled brain site it follows by definition that  $EC=N$  and  $Density=0$ .

Second, for two ECs there is no PDC necessary, i.e. for  $EC=X$  and  $EC=N$ .

The former EC expresses our inability to make any precise statement about the extension; thus, there is no precision to measure. The latter EC indicates that the respective brain site shows no label at all. The precision of this statement, however, is fully dependent on the precision by which the unlabelled brain site was identified and is thus already represented by the `PDC_BrainSiteIdentification`. Therefore, the `PDC_EC` for these two cases are represented by a dash (-).

Third, for the same reasons as above, there is no `PDC_Density` for  $Density=X$  or  $Density=0$ . Again, the `PDC_Density` is represented by a dash (-) for these cases.

### ***G. Precision of Description Codes (PDCs)***

A common problem when comparing several experimental findings between and even within articles is that the data can be described at very different levels of precision. Different authors usually follow different strategies for presenting their experimental results from tract tracing studies. To juxtapose two extremes of the spectrum, some authors precisely relate their findings of existent or absent label to carefully defined brain sites and provide additional figures with clearly delineated borders, clear names of brain sites, and clearly plotted label. Others may just refer to vaguely defined topographic regions and provide sketchy diagrams with approximate locations of label and without names and borders of brain sites. Generally, one can conclude that the less precise a description of data in the literature, the higher the variation between different "observers" of these data with respect to constructing representations of these data within a database. In other words, the objectivity (i.e. observer-independence) of any method for "observing" and interpreting data from the literature strongly depends on the precision of data description. These problems become very important as soon as contradictory statements from different articles are compared: which of the conflicting data should be assigned higher priority? Ideally, one would like to determine data validity by some kind of "truth index" for any given statement from the literature that is represented in CoCoMac, i.e. the probability that the described finding correctly describes the "true" connectivity of the Macaque brain. Decisions about conflicting reports could then be made on the basis of such indices. Unfortunately, there exist no clear criteria by which the different sources of variation described above could be assessed objectively, and with which their relative importance should be weighted. Within CoCoMac we refrain from defining such a "truth" index. Instead, we reformulate the above question for the "truth" of data in the literature. Applying criteria developed by theoretical psychology, "truth" of experimental data crucially depends on how these data are measured and represented by an observer. If the chosen method of observing and representing data fails to satisfy the fundamental measurement

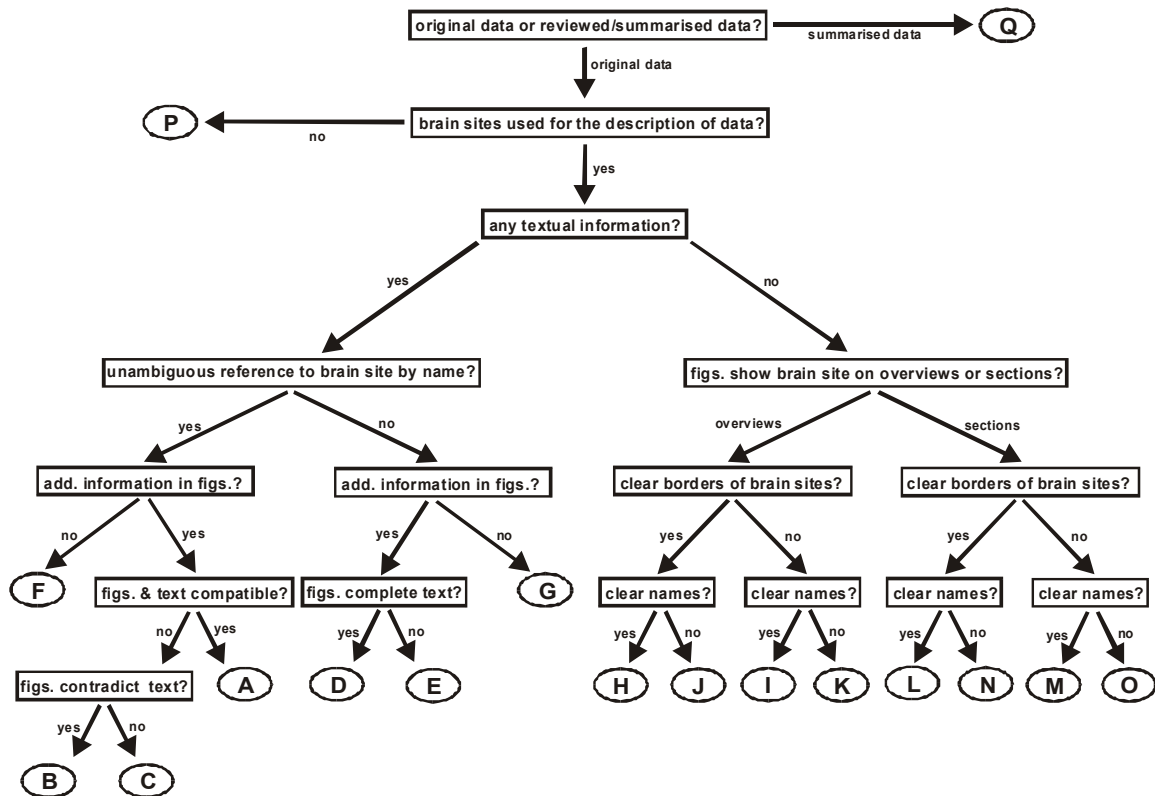
## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

criteria of objectivity, reliability, and validity, it becomes very difficult to distinguish between "correct" and "incorrect" data since errors / variation due to the experimental design can no longer be distinguished from errors / variance due to the method of observation and representation. Furthermore, these three criteria possess a hierarchical relationship among each other: high objectivity is a necessary (although not sufficient) condition for high reliability, which again is a necessary (although not sufficient) condition for high validity (Bortz & Döring 1995; Guilford 1954; Lienert & Ratz 1994). Therefore, it appears useful to assess primarily the objectivity of the data in the literature. As already mentioned above, this is feasible by estimating for any given datum the likelihood that individual data collators would have interpreted it differently and thus created non-equivalent representations of it in databases. In all subsequent data comparisons and analyses, whenever one needs to choose between contradictory statements from different articles, one would prefer that statement which was described most precisely and which is thus least likely to be confounded by errors of interpretation. As a scale for the descriptive precision of data, CoCoMac introduces a new coding scheme: "Precision of data Description Codes" (abbreviated as "PD Codes" or "PDCs"). The main principle of PDCs is to define a set of clear criteria concerning data representation within each data modality. For example, the identification of a cortical area that is labelled (or not labelled) by transported tracer substance can be accomplished by textual descriptions, tables, exact figures, diagrams, photographs, or by a combination thereof. The text may explicitly identify an area by its full name or acronym, or refer to it vaguely by topographic descriptions. Figures may or may not show clear areal borders, specific areal names, and the exact extent of tracer substance. Text and figures may confirm, complement, or contradict each other, and so on. Regarding each combination of such criteria as a specific case, one obtains a set of disjoint classes with different degrees of ambiguity (see figure below). It is possible to define these criteria very carefully, so that the overall coding is well operationalised and thus standardised between different observers.

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany



Hierarchical decision tree for PDCs concerning the identification of (un)labeled or injected BrainSites.

It should be noted that even within a given article, the precision of data description may vary considerably. For example, a particular article may very precisely identify the brain sites that were labelled after a particular injection, while providing rather ambiguous statements on the density or the laminar patterns of the label. Alternatively, the precision of data description may vary from finding to finding within the same class of items. Therefore, PD coding is not only applied once to each article contained by CoCoMac, but to every individual datum of the following modalities: (i) identification of injected and (un)labelled brain sites (PDC\_Site), (ii) extent of tracer substance or transported label within an identified area (PDC\_EC), (iii) density / strength of label within an identified area (PDC\_Density), (iv) laminar patterns of label within an identified area (PDC\_Laminae), (v) inter-map relations of two brain sites (PDC\_RC). On the whole, PDCs can be conceptualised as a systematic observation method covering five categories. These five categories are described in detail in paragraphs G.1 – G.5.

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

Generally, it should be noted that, although we label the individual cases by consecutive upper-case letters, the alphabetical order is not meant to imply any pre-defined hierarchy among the PDCs. Instead, before starting the analysis routines of CoCoMac, the user can define variable PDC hierarchies, depending on personal preferences. For example, one might assign a higher value to data based on textual descriptions than on data merely displayed by figures and thus rank the PDCs accordingly. If contradictions between text and figures should be met during data analysis, the text-based data would win over the figure-based data. Other users may have the opposite preference and thus choose a different PDC hierarchy. Therefore, whenever we speak of "better" or "higher" PDC, we do not define superiority or inferiority in absolute terms, but always relative to a given user-chosen hierarchy.

In summary, PDCs are thus a means of indirectly estimating the validity of data from the literature at the early stage of objectivity (defined as observer-independent representation) and are an important element for all algorithms used by CoCoMac to analyse and integrate its data according to the user's preferences.

The following sections describe the definitions of PDCs for the existing 5 categories and address a key question that guides the choice of a particular PDC for a given datum in the literature, i.e. whether information is unambiguous or not.

## G.1 Identification of brain sites

In the context of the following definitions, textual information is considered unambiguous only if the brain site is referenced directly by its name or acronym (e.g. "Label was found in area 46."). Reference to topographical landmarks is not considered unambiguous (e.g. "Label was found in the vicinity of the principal sulcus.").

Information provided by figures is only considered as unambiguous if the brain site is identified by clear names and/or borders.

### PDC\_BrainSite

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

Code	Info in Text	Clear text	Info in Figs	Correspondence	Contradiction	Completion	Figs: O/S	Figs: names	Figs: borders
<b>A</b>	+	+	+	+	-				
<b>B</b>	+	+	+	-	+				
<b>C</b>	+	+	+	-	-				
<b>D</b>	+	-	+			+			
<b>E</b>	+	-	+			-			
<b>F</b>	+	+	-						
<b>G</b>	+	-	-						
<b>H</b>	-		+				O	+	+
<b>I</b>	-		+				O	+	-
<b>J</b>	-		+				O	-	+
<b>K</b>	-		+				O	-	-
<b>L</b>	-		+				S	+	+
<b>M</b>	-		+				S	+	-
<b>N</b>	-		+				S	-	+
<b>O</b>	-		+				S	-	-
<b>P</b>	For this datum, the article does not refer to a delineated brain site by its name or acronym, but (i) uses broad topographic labels to describe its data OR (ii) displays data on figures which lack areal names and borders and for which none of the map(s) defined or adopted by the article is valid.								
<b>Q</b>	Information on area is from a review article.								

Case A: The (un)labelled area is named explicitly (i.e. referred to by its name) in the text/tables and identified with certainty. Additional figures explicitly support the text by showing present (or missing) label in areas defined by names and/or borders.

Case B: The (un)labelled area is named explicitly (i.e. referred to by its name) in the text/tables and identified with certainty. Additional figures contradict the text by showing present (or missing) label in the respective area defined by names and/or borders.

Case C: The (un)labelled area is named explicitly (i.e. referred to by its name) in the text/tables and identified with certainty. Additional figures contain information that neither supports nor contradicts the text, e.g. by showing label in brain regions without areal names and borders.

Case D: The (un)labelled area is referred to by the author in the text/tables, but its identification is not clear (i.e. the author does not denote the area directly by its name). However, additional figures achieve a clear identification by showing the area with a clear name and/or clear borders.

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

- Case E: The (un)labelled area is referred to by the author in the text/tables, but its identification is not clear (i.e. either the author does not denote the area directly by its name or explicitly expresses his uncertainty about correct identification). Additional figures do not help to achieve a clear identification as they lack clear areal names and borders.
- Case F: The (un)labelled area is named explicitly by the author in the text/tables by its name and identified unambiguously. Figures are not shown.
- Case G: The (un)labelled area is referred to by the author in the text/tables, but identified without certainty (that is, either the author himself mentions doubts about unequivocal identification of the area or he does not denote the area directly by its name). Figures are not shown.
- Case H: The (un)labelled area is not referred to by the text/tables. The figures provide an overview of the regional cortex where the area is situated and depict the (un)labelled area with clear areal names and clear areal borders.
- Case I: The (un)labelled area is not referred to by the text/tables. The figures provide an overview of the regional cortex where the area is situated and depict the (un)labelled area with clear areal names, but unclear/missing areal borders.
- Case J: The (un)labelled area is not referred to by the text/tables. The figures provide an overview of the regional cortex where the area is situated and depict the (un)labelled area with clear areal borders, but missing/unclear areal names.
- Case K: The (un)labelled area is not referred to by the text/tables. The figures provide an overview of the regional cortex where the area is situated and depict the (un)labelled area, but show neither clear areal names and nor clear areal borders, so that the label can be assigned to an area of the author's map only indirectly by its position and extension.

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

Case L: The (un)labelled area is not referred to by the text/tables. The figures show brain sections of the regional cortex where the area is situated and depict the (un)labelled area with clear areal names and clear areal borders.

Case M: The (un)labelled area is not referred to by the text/tables. The figures show brain sections of the regional cortex where the area is situated and depict the (un)labelled area with clear areal names, but unclear/missing areal borders.

Case N: The (un)labelled area is not referred to by the text/tables. The figures show brain sections of the regional cortex where the area is situated and depict the (un)labelled area with clear areal borders, but missing/unclear areal names.

Case O: The (un)labelled area is not referred to by the text/tables. The figures show brain sections of the regional cortex where the area is situated but show neither clear areal names and nor clear areal borders, so that the label can be assigned to an area of the author's map only indirectly by its position and extension.

Case P: For this datum, the article does not refer to a delineated brain site by its name or acronym, but (i) uses broad topographic labels to describe its data OR (ii) displays data on figures which lack areal names and borders and for which none of the map(s) defined or adopted by the article is valid.

Case Q: The information about the (un)labelled area is not from an original research report, but from a review article.

## G.2 Extension Codes

In the context of the following definitions, textual information is considered as unambiguous if the spatial extent of tracer substance within the brain site of interest is expressed clearly (e.g. "Only parts of area 46 showed labelled neurons.") or implied directly ("We found label in the dorsal part of area 46."). Information provided by figures is only considered unambiguous if clear borders delineate the respective brain site.

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

With respect to information on the EC, figures showing brain sections are problematic as they only represent a tiny fraction of an brain site in question, e.g. even in an brain site that is completely filled with label at small regular intervals, some sections may show missing or incomplete labelling. On the other hand, within a partially labelled brain site a given section might still show label throughout its entire extent.

⇒ **Therefore, information provided by sections is not suitable for confirming or refuting textual information on the extension of label** (i.e.  $PDC\_EC \notin \{A, B, D\}$  for this case). **Only if no textual information on the EC is given at all, information provided by sections may be used** (i.e.  $PDC\_EC \in \{L, M, N, O\}$ ).

### PDC\_EC

Code	Info in Text	Clear text	Info in Figs	Correspondence	Contradiction	Completion	Figs: O/S	Figs: Borders	Figs: label
A	+	+	+	+	-				
B	+	+	+	-	+				
C	+	+	+	-	-				
D	+	-	+			+			
E	+	-	+			-			
F	+	+	-						
G	+	-	-						
H	-		+				O	+	+
I	-		+				O	+	-
J	-		+				O	-	+
K	-		+				O	-	-
L	-		+				S	+	+
M	-		+				S	+	-
N	-		+				S	-	+
O	-		+				S	-	-

Case A: The extension of the label within an area is stated unambiguously by the text/tables (e.g. "Label was only found in the dorsal part of Walker's area 46."). Additional figures support the text by showing clearly indicated label in areas with clear borders. **NOTE:** for 2D areas, brain sections are, in principle, not able to provide information about the entire area and thus cannot confirm clear textual information about the EC.

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

Case B: The extension of the label within an area is stated unambiguously by the text/tables (e.g. "Label was only found in the dorsal part of Walker's area 46."). Additional figures, showing clearly indicated label in areas with clear borders, contradict the text. **NOTE:** for 2D areas, brain sections are, in principle, not able to provide information about the entire area and thus cannot refute textual information about the EC.

Case C: The extension of the label within an area is stated unambiguously by the text/tables (e.g. "Label was only found in the dorsal part of Walker's area 46."). Additional figures contain information that neither supports nor contradicts the text (e.g. by showing labelled brain regions, whose identification is not unambiguous due to lack of areal names and borders).

**NOTE:** This PDC\_EC also applies to those cases of 2D areas for which exists clear textual information and additional information provided by sections because the latter can, in principle, neither confirm nor contradict a given textual statement on the EC.

Case D: Information about the extension of label within an area is given by the text/tables only partially or ambiguously, but is complemented by additional figures showing clearly indicated label in areas with clear borders. **NOTE:** for 2D areas, brain sections are, in principle, not able to provide information about the entire area and thus cannot clarify ambiguous textual information about the EC.

Case E: Information about the extension of label within an area is given by the text/tables only partially or ambiguously. Additional figures show label in areas without clear borders and thus do not resolve uncertainty completely.

**NOTE:** This PDC\_EC also applies to those cases for which exists ambiguous textual information and additional information provided by sections because the latter can, in principle, neither confirm nor contradict a given textual statement on the EC.

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

Case F: The extension of the label within an area is stated unambiguously by the text/tables (e.g. "Label was only found in the dorsal part of Walker's area 46."). Additional figures are not shown.

Case G: Information about the extension of label within an area is given by the text/tables only partially or ambiguously. Additional figures are not shown.

Case H: The extension of the label within an area is not stated by the text/tables. The figures show a precise description of the label and areas with clear borders. They provide an overview of the regional cortex where the area is situated.

Case I: The extension of the label within an area is not stated by the text/tables. The figures show areas that have clear borders, but do not describe the label precisely. They provide an overview of the regional cortex where the area is situated.

Case J: The extension of the label within an area is not stated by the text/tables. The figures show a precise description of the label, but no clear areal borders. They provide an overview of the regional cortex where the area is situated.

Case K: The extension of the label within an area is not stated by the text/tables. The figures neither show a precise description of the label nor areas with clear borders. They provide an overview of the regional cortex where the area is situated.

Case L: The extension of the label within an area is not stated by the text/tables. The figures show a precise description of the label and areas with clear borders. They show brain sections of the regional cortex where the area is situated.

**NOTE:** For 2D areas, this case is only applicable if neither textual information on the EC nor figures with overviews are provided.

Case M: The extension of the label within an area is not stated by the text/tables. The figures show areas that have clear borders, but do not describe the label precisely. They show brain sections of the regional cortex where the area is situated.

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

**NOTE:** For 2D areas, this case is only applicable if neither textual information on the EC nor figures with overviews are provided.

Case N: The extension of the label within an area is not stated by the text/tables. The figures show a precise description of the label, but no clear areal borders. They show brain sections of the regional cortex where the area is situated.

**NOTE:** For 2D areas, this case is only applicable if neither textual information on the EC nor figures with overviews are provided.

Case O: The extension of the label within an area is not stated by the text/tables. The figures neither show a precise description of the label nor areas with clear borders. They show brain sections of the regional cortex where the area is situated.

**NOTE:** For 2D areas, this case is only applicable if neither textual information on the EC nor figures with overviews are provided.

## G.3 Density of label

In the context of the following definitions, textual information is considered unambiguous only if one of the following key words is used:

weak, sparse, light labelling → density 1

moderate labelling → density 2

strong, dense, heavy labelling → density 3

Descriptions as "...a large number of labelled neurons..." or "...only a few scattered neurons were found..." are not treated as unambiguous textual information.

Information provided by figures is only considered unambiguous if symbols are used that clearly assign the label within an brain site to one of the three classes above (e.g. if triangles denote weak, circles moderate, and squares strong labelling). Information by figures that require interpretation of the density of plotted neurons is not considered unambiguous (PDC\_Density = I must be assigned when interpreting such figures).

### PDC\_Density

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

Code	Info in Text	Clear text	Info in Figs.	Correspondence	Contradiction	Completion	Clear figs.	Rule
A	+	+	+	+	-			
B	+	+	+	-	+			
C	+	+	+	-	-			
D	+	-	+			+		
E	+	-	+			-		
F	+	+	-					
G	+	-	-					
H	-		+				+	
I	-		+				-	
J								+

Case A: The density of labelling is stated by the text/tables in one of the following terms: dense/heavy, moderate/medium, sparse/light. Information provided by figures supports the text.

Case B: The density of labelling is stated by the text/tables in one of the following terms: dense/heavy, moderate/medium, sparse/light. Information provided by figures is contradictory to that of the text.

Case C: The density of labelling is stated by the text/tables in one of the following terms: dense/heavy, moderate/medium, sparse/light. Information provided by figures neither supports nor contradicts the text.

Case D: The density of labelling is described by the text/tables in an imprecise or ambiguous way (e.g. BP89: + = light-moderate labelling, ++ = moderate-heavy labelling or descriptions of sparse labelling like "...a few scattered neurons..."), but information provided by figures complements the text.

Case E: The density of labelling is described by the text/tables in an imprecise or ambiguous way. Information provided by figures does not help to achieve certainty.

Case F: The density of labelling is stated by the text/tables unambiguously in one of the following terms: dense/heavy, moderate/medium, sparse/light. Figures contain no information about labelling density.

Case G: The density of labelling is described by the text/tables in a way that requires interpretation. Figures contain no information about labelling density.

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

Case H: The density of labelling is not described by the text/tables, but the figures contain unambiguous information (e.g. by assigning each of the 3 density classes to a particular symbol).

Case I: The density of labelling is not described by the text/tables. Figures contain information that is ambiguous or needs to be interpreted (e.g. figures that mark each individual labelled neuron by a dot and thus reflect the relative density of the labelling).

Case J: The density of labelling is not described with reference to a particular projection, but by a general rule (e.g. "All projections from area MII to prefrontal areas showed dense labelling.")

### G.4 Laminar patterns

In the context of the following definitions, textual information is considered unambiguous if it is directly expressed which laminae contain tracer substance (e.g. "Labelled neurons were found in laminae III and V." or "Label extended through all cortical layers.") or if clear references to supragranular or infragranular layers are used (e.g. "Labelled neurons were found in both supra- and infragranular layers, but spared layer IV."). In case of supragranular labelling, however, it is usually not stated clearly if layer I is included or not. Therefore, layer I should be coded as "?" in these cases.

Information provided by figures is considered unambiguous if

- on overviews: symbols are used that clearly assign the label within an brain site to a specific type of laminar pattern (see figures of GP83 for an example)
- on sections: clear borders between the laminae are shown so that plotted label can be assigned to the laminae.

Information by figures that show the location of labelled neurons on sections without clear borders require interpretation and is thus not considered unambiguous.

**REMEMBER:** laminae are represented differently!

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

isocortical laminae (except V1) → representation by the standard 6-tupel  
as a property of a labelled brain site

laminae of area V1

allocortical laminae → represented as independent brain sites

subcortical laminae

### PDC\_Laminae

Code	Info in Text	Clear text	Info in Figs.	Correspondence	Contradiction	Completion	Clear figs.	Rule
<b>A</b>	+	+	+	+	-			
<b>B</b>	+	+	+	-	+			
<b>C</b>	+	+	+	-	-			
<b>D</b>	+	-	+			+		
<b>E</b>	+	-	+			-		
<b>F</b>	+	+	-					
<b>G</b>	+	-	-					
<b>H</b>	-		+				+	
<b>I</b>	-		+				-	
<b>J</b>								+

Case A: The laminar pattern of the projection is stated by the text/tables in an unambiguous way (e.g. "The projection terminates in layers III, V and VI only."). Additional figures support the text by showing label in delineated laminae or by assigning specific symbols to particular laminar patterns.

Case B: The laminar pattern of the projection is stated by the text/tables in an unambiguous way. Additional figures clearly contradict the text by showing label in delineated laminae or by assigning specific symbols to particular laminar patterns.

Case C: The laminar pattern of the projection is stated by the text/tables in an unambiguous way. Additional figures contain ambiguous information that neither clearly supports nor clearly contradicts the text (e.g. by showing brain sections with labelled cells but without clear delineation of the laminae).

Case D: Information about the laminar pattern of the projection is given by the text/tables in a way that requires interpretation, but is complemented

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

by additional figures that show label in delineated laminae or by assigning specific symbols to particular laminar patterns.

Case E: Information about the laminar pattern of the projection is given by the text/tables in a way that requires interpretation. Additional figures do not clearly complement the text (e.g. by showing brain sections with labelled cells but without clear delineation of the laminae).

Case F: The laminar pattern of the projection is stated by the text/tables in an unambiguous way (e.g. "The projection terminates in layers III, V and VI only."). Figures contain no information about laminar patterns.

Case G: The laminar pattern of the projection is stated by the text/tables in a way that requires interpretation (e.g. "Labelled cells were found mainly in the supragranular layers."). Figures contain no information about laminar patterns.

Case H: The laminar pattern of the projection is stated by the figures in an unambiguous way by showing label in delineated laminae or by assigning specific symbols to particular laminar patterns. Text/tables contain no information about laminar patterns.

Case I: The laminar pattern of the projection is stated by the figures in a way that requires interpretation. Text/tables contain no information about laminar patterns.

Case J: The only source of information about the laminar pattern of the projection is a general rule described by the text (e.g. "Projections from area 46 to all other prefrontal areas terminate in layers III and V.>").

### G.5 Relations between BrainMaps

In the context of the following definitions, textual information is considered unambiguous only if the relation is directly expressed (e.g. "Our area X is identical with Walker's (1940) area 14." or "Area X includes areas V4 and V4t of Felleman & Van Essen 1991."). It is not considered unambiguous, however, if the text expresses uncertainty about the relation (e.g. "Our area X seems to

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

be identical with Walker's (1940) area 14." or "Judged by its location, area X appears to include areas V4 and V4t of Felleman & Van Essen 1991."). It is not considered unambiguous either if the text expresses a relation without specifying it, e.g. by simply juxtaposing two areas: "...particularly interesting was a finding in area FB of von Bonin & Bailey (area 6 of Brodmann 1909)...". Although such juxtaposition suggests identity, it may reflect any of the 4 possible relations. In this example, the relation would be RC (BB47-FB, B09-6) = S!

Information provided by figures is only considered unambiguous if the figures contain clear borders and either show the two brain sites projected onto the same brain or display them on two juxtaposed figures of the same (standardized) brain. Juxtaposition of two brain sites on different (individual) brains, is still helpful to determine relations but is not considered unambiguous.

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

## PDC\_Relation

Code	Info in text	Clear text	Info in figs	Correspondence	Contradiction	Completion	Clear ref.	Clear figs	Other cases
A	+	+	+	+	-				
B	+	+	+	-	+				
C	+	+	+	-	-				
D	+	-	+			+			
E	+	-	+			-			
F	+	-	+			-	+		
G	+	-	+			-	-		
H	+	+	-						
I	+	-	-						
J	+	-	-				+		
K	+	-	-				-		
L	-		+					+	
M	-		+					-	
N	-		+				+	-	
O	-		+				-	-	
P	-		-						Assumption
Q	-		-						GM
R	-		-						AM

Case A: The relation of a brain site in the source map to another brain site in the target map is stated by the text/tables in an unambiguous way. Additional figures support the text by either showing the two areas projected onto the same brain or by showing them on two juxtaposed figures of the same (standardized) brain.

Case B: The relation of a brain site in the source map to another brain site in the target map is stated by the text/tables in an unambiguous way. Additional figures contradict the text by either showing the two areas projected onto the same brain or by showing them on two juxtaposed figures of the same (standardized) brain.

Case C: The relation of a brain site in the source map to another brain site in the target map is stated by the text/tables in an unambiguous way. Additional figures contain ambiguous information that neither supports nor contradicts the text by showing the two maps on two different and not directly comparable brains.

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

- Case D: Information about the relation of a brain site in the source map to another brain site in the target map is given by the text/tables in a way that requires interpretation, but is complemented by additional figures which either show the two areas projected onto the same brain or show them on two juxtaposed figures of the same (standardized) brain. The figures thus resolve uncertainty.
- Case E: Information about the relation of a brain site in the source map to another brain site in the target map is given by the text/tables in a way that requires interpretation. Additional figures do neither show the two areas projected onto the same brain nor do they show them on two juxtaposed figures of the same (standardized) brain. Thus, they do not manage to resolve the uncertainty completely.
- Case F: Information about the relation of a brain site in the source map to another brain site in the target map is given by the text/tables in a way that requires interpretation. Additional figures do neither show the two areas projected onto the same brain nor do they show them on two juxtaposed figures of the same (standardized) brain. Thus, the figures do not manage to resolve uncertainty completely. However, other articles which are being referred to for this relation eliminate the ambiguity by either clear textual statements or clear figures.
- Case G: Information about the relation of a brain site in the source map to another brain site in the target map is given by the text/tables in a way that requires interpretation. Additional figures do neither show the two areas projected onto the same brain nor do they show them on two juxtaposed figures of the same (standardized) brain. Thus, the figures do not manage to resolve uncertainty completely. Other articles which are being referred to for this relation do not eliminate the ambiguity either.
- Case H: The relation of a brain site in the source map to another brain site in the target map is stated by the text/tables in an unambiguous way, whilst figures contain no information. This case also includes

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

relations between areas from companion articles (i.e. if author X publishes two articles in the same issue of a given journal, it is assumed that the areas used in the two articles are identical - unless they are explicitly stated to be different).

Case I: The relation of a brain site in the source map to another brain site in the target map is either stated by the text/tables in a way that requires interpretation or is explicitly mentioned by the author to be uncertain, whilst figures contain no information. This case includes statements where a given area is simultaneously related to several brain maps (e.g. "Area X is situated at the dorsal bank of the principal sulcus (ref1, ref2, ref3...)." ) or where two areas of different areas are simply juxtaposed without directly mentioning their relation (e.g. "Area 19 of Brodmann (OA of Bonin & Bailey)...").

Case J: The relation of a brain site in the source map to another brain site in the target map is either stated by the text/tables in a way that requires interpretation or is explicitly mentioned by the author to be uncertain, whilst figures contain no information. However, other articles which are being referred to for this relation eliminate the ambiguity by either clear textual statements or clear figures.

Case K: The relation of a brain site in the source map to another brain site in the target map is either stated by the text/tables in a way that requires interpretation or is explicitly mentioned by the author to be uncertain, whilst figures contain no information. Other articles which are being referred to for this relation do not eliminate the ambiguity either.

Case L: The relation of a brain site in the source map to another brain site in the target map is stated by the figures in an unambiguous way, i.e. they either show the two areas projected onto the same brain or show them on two juxtaposed figures of the same (standardized) brain. Text/tables contain no information.

Case M: The relation of a brain site in the source map to another brain site in the target map is stated by the figures in a way that requires

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

interpretation, i.e. they show the two maps on two different and not directly comparable brains. Text/tables contain no information.

Case N: The relation of a brain site in the source map to another brain site in the target map is stated by the figures in a way that requires interpretation, i.e. they show the two maps on two different and not directly comparable brains. Text/tables contain no information.

However, other articles which are being referred to resolve the uncertainty by either textual or figural information.

Case O: The relation of a brain site in the source map to another brain site in the target map is stated by the figures in a way that requires interpretation, i.e. they show the two maps on two different and not directly comparable brains. Text/tables contain no information. Other articles which are being referred to do not manage to resolve the uncertainty either.

Case P: As there is no direct information available, the relation of a brain site in the source map to another brain site in the target map can only be assumed because of its name (e.g. if an area is designated by a term that is commonly linked to a specific brain map, like “area 17” which usually refers to the maps of Brodmann 1905, 1909) or by comparing the maps for gross anatomical landmarks.

Case Q: This code is reserved for relations that contain BrainSites of the General Map (GM). This concerns both (i) relations between BrainSites from original maps (OM) and GM and (ii) relations between BrainSites in GM only.

Case R: This code is reserved for relations that contain BrainSites of the Acronym Map (AM). Note that these relations are created automatically ONLY and are not to be entered manually.

## ***H. Guidelines for proofreading previously entered data***

### **H.1 General remarks**

When proofreading the representation of a previously entered article in CoCoMac, pay particular attention to the points outlined below. Also, it is strongly recommended to perform proofreading in exactly the same order in which data are entered, i.e. along the hierarchical sequence amongst the tables: (1) bibliographic data, (2) mapping data, (3) experimental data.

Red entries denote fields of particular importance. These fields may be crucial to correct interpretation of the data (e.g. hemisphere, other sites affected by an injection, etc.). Alternatively, they were introduced while CoCoMac was already under construction (e.g. SiteDef\_Type in LabelledSites\_Data), or their regulations were changed at some point (e.g. interpretation of density from figures). These fields must be checked particularly carefully.

### **H.2. Bibliographic data**

#### **H.2.1 Literature**

1. Are the fields set correctly that concern the status of the entered article (i.e. CorticalConnectivity, SubcorticalConnectivity, MappingData, PhysicalCopy)?
2. Is there a hyperlink to the PubMed abstract (if available)?
3. Check that the exact reference is entered into the appropriate table (i.e. Literature\_JournalArticles, Literature\_Books, Literature\_BookChapters).

#### **H.2.2 Literature\_LinkTable**

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

1. Are all authors entered?
2. Is the order of the authors' list correct?

## H.3 Mapping data

### H.3.1 BrainMaps

1. Are the fields `Delineation_BrainSites` and `UnspecificAdoption_BrainSites` set correctly?

### H.3.2 BrainMap\_BrainSiteAcronyms

For each acronym used by the article:

1. Does each acronym also have a full name ? Make sure that even seemingly trivial abbreviations are accompanied by a full name. For example, "46" should be explained by "prefrontal area 46".

### H.3.3 BrainMaps\_BrainSites

For each brain site defined or adopted by the article:

1. Does `SiteDef_Type` have a value for each brain site?
2. Is the `SiteClass` defined correctly?

### H.3.4 BrainMaps\_Methods

1. Are the methods represented by which the article defines its `BrainSites`?

### H.3.5 InterMapRelations

1. Check the right order in which the brain sites form the relation (this will soon be done algorithmically).

### H.3.6 InterMapRelations\_References

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

1. Check PDC\_Relation.
2. Are the references accompanied by a citation? If not, add it.
3. Check the Citation for comprehensiveness, completeness, and typos.

## H.4 Experimental data

### H.4.1 Injections

1. Check PDC\_Site.
2. Check whether the SiteType is correctly chosen and whether the EC is correct.
3. Check PDC\_EC (remember the caveat about PDC\_EC's if only sections are used to illustrate the results:  $PDC\_EC \notin \{A, B, D\}$  but  $PDC\_EC \in \{L, M, N, O\}$  – see manual, section F.2).
4. Check the Citation for comprehensiveness, completeness, and typos.
5. Check Hemisphere and AffctedNeighbours.
6. Check whether MethodicalProblems is set correctly. (remember that there are three main reasons to tick this checkbox: (i) infiltration of white matter by tracer substance, (ii) lesions of the traced brain due to the operation, (iii) more than one injection with identical tracers into more than one brain site (e.g. bihemispheric injections of the same tracer).

### H.4.2 Injections\_Laminae

1. Check whether the area that this laminar pattern is described for is really isocortical. Remember that if it is allocortical, its laminae must be represented as individual BrainSites.
2. Check PDC\_Laminae.

### H.4.3 LabelledSites\_Data

1. Check PDC\_Site.

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

2. Check whether the SiteType is correctly chosen and whether the EC is correct.
3. Check PDC\_EC (remember the caveat about PDC\_EC's if only sections are used to illustrate the results:  $\text{PDC\_EC} \notin \{A, B, D\}$  but  $\text{PDC\_EC} \in \{L, M, N, O\}$  – see manual, section F.2).
4. Check density and PDC\_Density (remember to evaluate also figures that indirectly indicate density by plotting individual neurons and assign  $\text{PDC\_Density}=1$ ).
5. Check whether QuantitativeData is set correctly.

### H.4.4 LabelledSites\_Data\_Laminae

1. Check whether the area that this laminar pattern is described for is really isocortical. Remember that if it is allocortical, its laminae must be represented as individual brain sites.
2. Check PDC\_Laminae.

### H.4.5 LabelledSites\_Descriptions

1. Check Hemisphere and Terminal\_vs\_Soma.

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

## *I. List of abbreviations and technical terms*

- **AM:** Acronym Map, a set of BrainSites that are algorithmically created to allow for relations between BrainSites with identical acronyms and likely, but non-documented identity → C.2.1
- **BrainMap:** a set of BrainSites, i.e. microstructurally and/or functionally defined topographical units (defined in a “brain space”) as presented by a particular article → C.1.1
- **BrainSite:** see BrainMaps
- **Delineation:** one of three possible cases in what way a given BrainSite can be defined in a particular article (see Specific Adoption, Unspecific Adoption); this case refers to BrainSites whose definitions have been explicitly stated by the article, either by a textual description or a graphical illustration → C.4.8
- **ECs:** Extension Codes, a set of codes that describe the spatial distribution of information on injected or transported tracer substance within a BrainSite.
- **GM:** General Map, a set of predefined, hierarchically arranged general BrainSites which allow to link BrainSites with different names, but identical meaning → C.2.1.
- **InterMapRelation:** a logical relation (denoted by an RC) between two BrainSites; often simply referred to as “relation”.
- **OM:** (i) set of all BrainMaps defined by original articles; (ii) one specific BrainMap defined by an original article
- **PDCs:** Precision of Description Codes, a general term form 5 different sets of codes that rate the precision by which experimental data is described in the literature.

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

- **RCs:** Relation Codes, a set of codes that describe the logical relationship between BrainSites from different BrainMaps.
- **SiteClass:** classification scheme that defines which kind of “module” a given BrainSite represents within the organizational hierarchy of the brain (i.e. whether it is a cortical or subcortical structure, an area, a lamina, a column, a single neuron, etc.) → C.3.
- **SiteType:** indicates whether a given BrainSite is used as a 2D or 3D “storage unit” for the description of connectivity data (i.e. whether any ECs that refer to this BrainSite reflect the extension of information within a 2D or 3D space) → C.3.
- **Specific Adoption:** one of three possible cases in what way a given BrainSite can be defined in a particular article (see Delineation, Unspecific Adoption); this case refers to BrainSites whose definitions have been explicitly stated to be adopted from another article → C.4.8
- **Unspecific Adoption:** one of three possible cases in what way a given BrainSite can be defined in a particular article (see Delineation, Specific Adoption); this case refers to BrainSites that have been adopted but the exact source of their definitions remains unclear → C.4.8

# Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

## History of changes

03-09-22	Adapted the rule for dealing with articles that have identical IDs in section B.1.	KES
03-01-20	Rephrased description of “Concentration” field in the Injections table.	KES
02-09-01	Added section C.2 on AM and GM.	KES
02-09-01	Added section I (list of abbreviations and technical terms).	KES
02-08-11	Included a general paragraph on PDCs in section G and added the tables for each PDC class to the following chapters G.1 – G.5.	KES
02-08-11	Included table of permissible RC/SiteClass constellations for InterMapRelations in section C.3.9	KES
02-08-11	Updated definition of SiteType “Band” in section C.3.2	KES
01-11-10	Updated information on administration tables (E.2., E.3)	KES
01-11-10	Added information on new Administration_Proofreading tables (B.6).	KES
01-11-10	Added information on new Status-fields in table Literature (B.1)	KES
01-09-09	Added warning not to define relations across more than one level (C.3.9).	KES
01-09-09	Updated information on Methods_Animals and Methods_Animals_Details (D.1.2, D.1.3).	KES
01-09-09	Changed GenSiteTypes and SpecSiteTypes to SiteClasses and SiteTypes throughout the manual.	KES
01-07-28	Added table of contents.	KES

## Computational Systems Neuroscience Group (CSN)

C. & O. Vogt Brain Research Institute, Düsseldorf / Germany

01-07-28	Added information about the new concept of GenSiteTypes and SpecSiteTypes to C.2. New tables are described at C.3.1 and C.3.2; new fields in experiment-related tables are described at D.1.4, D.1.5, D.1.8.	KES
01-07-28	Added indices to all tables.	KES
01-07-05	Added general introduction to BrainSiteTypes (C.3)	KES